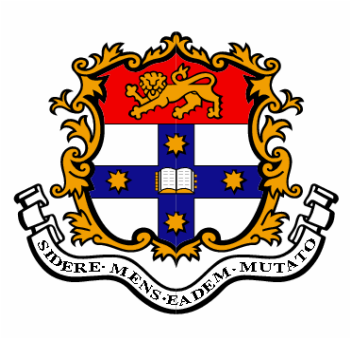
A SCRUTABLE ADAPTIVE HYPERTEXT



A Thesis submitted in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in the School of Information Technologies at
The University of Sydney

Marek Czarkowski

March 2006

3/6/2006 9:50:00 AM

Abstract

Fuelled by the popularity and uptake of the World Wide Web since the 1990s, many researchers and commercial vendors have focussed on Adaptive Hypermedia Systems as an effective mechanism for disseminating personalised information and services. Such systems store information about the user, such as their goals, interests and background, and use this to provide a personalised response to the user. This technology has been applied to a number of contexts such as education systems, e-commerce applications, information search and retrieval systems.

As an increasing number of systems collect and store personal information about their users to provide a personalised service, legislation around the world increasingly requires that users have access to view and modify their personal data. The spirit of such legislation is that the user should be able to understand how personal information about them is used. There literature has reported benefits of allowing users to access and understand data collected about them, particularly in the context of supporting learning through reflection. Although researchers have experimented with open user models, typically the personalisation is inscrutable: the user has little or no visibility in to the adaptation *process*. When the adaptation produces unexpected results, the user may be left confused with no mechanism for understanding why the system did what it did or how to correct it.

This thesis is the next step, giving users the ability to see what has been personalised and why. In the context of personalised hypermedia, this thesis describes the first research to go beyond open, or even scrutable user models; it makes the *adaptivity* and *associated processes* open to the user and controllable. The novelty of this work is that a user of an adaptive hypertext system might ask *How was this page personalised to me?* and is able to see just how their user model affected what they saw in the hypertext document. With an understanding of the personalisation process and the ability to control it, the user is able to steer the personalisation to suit their changing needs, and help improve the accuracy of the user model.

Developing an interface to support the scrutinisation of an adaptive hypertext is difficult. Users may not scrutinise often as it is a distraction from their main task. But when users need to scrutinise, perhaps to correct a system misconception, they need to easily find and access the scrutinisation tools. Ideally, the tools should not require any training and users should be able to use them effectively without prior experience or if have not used them for a long time, since this is how users are likely to scrutinise in practice.

The contributions of thesis are: (1) SASY/ATML, a domain independent, reusable framework for creation and delivery of scrutable adaptive hypertext; (2)a toolkit of graphical tools that allow the user to scrutinise, or inspect and understand what personalisation occurred and control it; (3) evaluation of

the scrutinisation tools and (4) a set of guidelines for providing support for the scrutinisation of an adaptive hypertext through the exploration of several forms of scrutinisation tools.

Keywords: Open user modelling, adaptive hypertext, scrutinisation.

Acknowledgements

…

Table of Contents

List of Tables

List of Figures

List of Related Publications

The following is a list of technical reports, conference and workshop papers and posters, that describe work that is part of this research and lead shape, and provided an avenue for peer feedback of, the research submitted in this thesis.

1998

[1] Czarkowski, M. (1998) ATML, Tutor and ADAPT: an adaptive hypertext teaching system. Technical Report, TR 519. University of Sydney Australia. ISBN 1 86487 044 3.

- *Technical report describing Tutor 1.*

2000

[2] Czarkowski, M. and Kay, J. (2000) Bringing scrutability to adaptive hypertext teaching. In: G. Gauthier, C. Frasson and K. VanLehn (eds.) Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol. 1839, (Proceedings of 5th International Conference on Intelligent Tutoring Systems (ITS'2000), Montreal, Canada, June 21-23, 2000) Berlin: Springer-Verlag, pp. 421-432.

- *Conference paper describing Tutor 1 and quantitative evaluation.*

2001

[3] Czarkowski, M., Kay, J. (2001) Tutor: support for scrutably personalised documents. ADCS 2001 Proceedings of the Sixth Australasian Document Computing Symposium, Volume 1, pp.29-36, Vercoustre, AM & Kay, J, Basse. 2001. ISBN 1-86487-440-6.

- *Conference paper describing Tutor 2.*

2002

[4] Czarkowski, M. and Kay, J. (2002) A scrutable adaptive hypertext. In: P. De Bra, P. Brusilovsky and R. Conejo (eds.) Proceedings of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002) Proceedings, Málaga, Spain, May 29-31, 2002, pp. 384-387.

- *Conference paper describing Tutor 2.*

2003

[5] Czarkowski, M., Kay, J. (2003a). How to give the user a sense of control over the personalization of adaptive hypertext? Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, User Modeling 2003 Session, pp. 121-132.

- *Workshop paper describing Tutor 3 and qualitative evaluation.*

[6] Czarkowski, M., Kay, J. (2003b). Challenges of Scrutable Adaptivity. Poster in: AIED 2003: proceedings from the 11th International Conference in Artificial Intelligence in Education; Sydney, Australia, July 20-24 Ulrich Hoppe, Felisa Verdejo, Judy Kay. IOS Press. pp. 404 - 407.

- *Poster paper describing Tutor 3 and qualitative evaluation.*

2005

[7] Czarkowski, M., Kay, J. and Potts, S. (2005a) Web Framework for Scrutable Adaptation. In: Proceedings of Workshop on Learner Modelling for Reflection, to Support Learner Control, Metacognition and Improved Communication between Teachers and Learners at 12th International Conference on Artificial Intelligence in Education, AIED'2005, Amsterdam, July 19, 2005, IOS Press, pp. 11-18.

- *Workshop paper describing Cell-Tutor and qualitative evaluation.*

[8] Czarkowski, M., Kay, J. and Potts, S. (2005b) Scrutability as a core interface
Poster in Proceedings of 12th International Conference on Artificial Intelligence in Education, AIED'2005, Amsterdam, July 19, 2005, IOS Press, pp. 783-786.

- *Workshop paper describing Cell-Tutor and qualitative evaluation.*

[9] Czarkowski, M. (2005c) Evaluating Scrutable Adaptive Hypertext. In proceedings from the 10th International conference on User Modelling, Edinburgh, UK, 24th-29th July 2005. Available online at http://www.easy-hub.org/hub/workshops

- *Workshop paper describing SASY 1 qualitative evaluation design.*

2006

[10] Czarkowski, M., Kay, J. (2006) Giving learners a real sense of control over adaptivity, even if they are not quite ready for it yet, Chen S and G Magoulas (eds), *Advances in Web-based Education: Personalized Learning Environments*, IDEA Publishing, Chapter 5, pp. 93-105.

- *Book chapter describing Tutor 3 and qualitative evaluation.*

Chapter 1 Introduction

This thesis is an exploration of how to provide effective support for user scrutiny of personalisation of an adaptive hypertext.

Adaptive systems typically store information about the user, such as their goals, interests and background, in what is called the user model. The information in the user model is used to provide a personalised response to the user. Adaptive hypertext systems, a class of adaptive systems, process a user's user model to create a hypertext document that is tailored to that individual.

Adaptive systems have been applied to a variety of domains including education, information filtering and search, adaptive help, and retrieval and e-commerce (particularly online shopping). The adaptivity and individualised response may enhance the user's experience or help them achieve their goal, for example, finding information more quickly.

As an increasing number of systems collect and store personal information about their users to provide a personalised service, legislation around the world increasingly requires that users have access to view and modify their personal data. The spirit of such legislation is that the user should also be able to understand how personal information about them is used. In parallel, the literature has reported other benefits of allowing users to access and understand data collected about them, particularly in the context of supporting learning through reflection.

Driven by this, researchers have experimented with open, inspectable and even scrutable user models. Their research explored how the user of an adaptive system might access, view and perhaps modify their user model. However, they have not explored how to make the process driving the adaptation, open, understandable and controllable by the user.

In the context of personalised hypermedia, this thesis goes beyond open, or even scrutable user models; it makes the *adaptivity* and *associated processes* open to the user and controllable. The contribution of this work is that a user of an adaptive hypertext system might ask *How was this page personalised to me?* and is able to see just how their user model affected what they saw in the hypertext document. With an understanding of the personalisation process and the ability to control it, the user is able to steer the personalisation to suit their changing needs.

The following fictional, futuristic scenario illustrates a situation where a person may wish to scrutinise the personalisation of an adaptive system:

While Marek is at work, Hal, his Intelligent Internet fridge, determines which household items have been consumed and places an online purchase order for the household groceries. When Marek returns home that day, he commences packing away the groceries that have been delivered. He is upset to discover a bottle of Diet Cola drink has been delivered instead of regular Cola. Marek interrogates Hal about the order:

Marek: Hal, why did you order Diet Coke instead or regular Coke?

Hal: Marek, you said you were starting a diet this week. I emailed Dr. Clarke, your nutritionist, who provided me with a list of banned food items. Dr. Clarke suggested you try Diet Coke as an alternative to regular Coke as it has less kilojoules.

Marek: Oh…..I see. What other "alternative" items did you order on account of my diet?

Hal: Skim milk, poly-unsaturated margarine and wine instead of beer.

Marek: OK, that stuff can stay but the Diet Coke has to go back. Please order regular Coke.

Scenario 1 – Scrutinising personalisation in the household of the future.

There are a number of interesting elements in this scenario. For one thing, Hal - the Internet fridge, is significantly more intelligent than a regular refrigerator. It knows personal things about Marek, for example, his goal to lose weight. Also, to help Marek achieve this goal, it personalised the grocery shopping list, substituting high-fat items with low-fat items. At its core, this fridge is an adaptive system that maintains a model of it users and processes this information to personalise its service to them.

Of particular interest, from the perspective of this thesis, Hal is held accountable for the personalised actions and decisions it makes on behalf of Marek. Marek was able to interrogate Hal to (1) understand Hal ordered Diet Coke to help Marek achieve his weight loss goal (2) find out what other items Hal had substituted to suit the new diet requirements. Furthermore, Marek was able to correct Hal to have it order regular Coke. By doing so, Marek is helping Hal refine its beliefs about him. This makes Marek's user model more accurate and in turn improves the personalised service Marek receives.

In other words, Marek was able to *scrutinise* Hal's personalised actions, as well as the user model.

As there are a number of challenges before Scenario 1 can become a reality, this thesis focuses on supporting scrutability of personalisation in a certain class of adaptive systems: adaptive hypermedia systems. Fuelled by the popularity and uptake of the World Wide Web since the 1990s, many

researchers and commercial vendors have focussed on Adaptive Hypermedia Systems as an effective mechanism for disseminating personalised information and services to large audiences through the World Wide Web (Brusilovsky, 1996; 2001; Kobsa & Josef, 2000; De Bra et al., 2004; 2002). This technology has been used in a number of contexts:

- Education systems, for example, ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001);
- E-commerce applications, for example, online retailers such as www.amazon.com;
- Information retrieval systems, for example the www.google.com search engine;
- Information systems and other applications.

Although the use of adaptive hypermedia is widespread, there is a lack of support for scrutinisation. The user has little or no visibility into the adaptation process. When the adaptation produces unexpected results, the user may be left confused with no mechanism for understanding why the system did what it did or how to correct it. Typically, the personalisation is inscrutable.

This thesis contributes a framework for creation and delivery of adaptive hypertext that the user can scrutinise, or inspect, to understand what personalisation occurred and how to control it. The main components of this framework are SASY 4 (Scrutable Adaptive hypertext System) and ATML 2 (Adaptive Tutorial Mark-up Language). ATML is a schema for the definition of scrutable adaptive hypertext content. SASY is an end-user web based system that processes ATML documents to produce a personalised hypertext document to the user based on their user model. SASY provides a set of interactive, graphical tools that allow the user to scrutinise how a hypertext document was personalised to them. The relationship between these components is illustrated in Figure 1.1. SASY 4 and ATML 2 are the end product of a series earlier systems developed and studied as part of this research: Tutor 1, Tutor 2, Tutor 3, Cell-Tutor, SASY 1, SASY 2 and SASY 3. The evolution of the scrutinisation tools is described in the thesis.

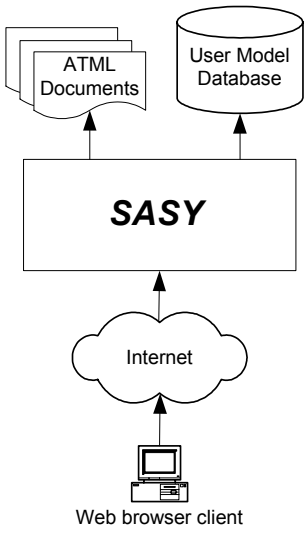


Figure 1.1 - Scrutable Adaptive Hypertext framework - SASY and ATML.

1.1 Definitions of key terms

Hypertext and Hypermedia

Hypermedia was first proposed by Vannevar Bush in 1945. Bush provided his vision of browsing the Web of linked information, which included the ability for readers to add their own information to add to the growing Web[1]. Theodor Holm Nelson first coined the terms hypertext and hypermedia (Nelson,1965):

"By `hypertext' I mean non-sequential writing--text that branches and allows choice to the reader, best read at an interactive screen. As popularly conceived, this is a series of text chunks connected by links which offer the reader different pathways."

Nelson also invented Xanadu, an architecture similar, but superior, to that of the World Wide Web. Jakob Nielsen (1990) added: *"Hypertext should also make users feel that they can move freely through the information according to their own needs."*

Hypertext is textual data that is linked across multiple documents or locations[2]. Today, the most popular form of hypertext is HyperText Mark-up Language (HTML), a standard developed and maintained by the World Wide Web Consortium (W3C)[3], and it is generally accessed through web browsers that request documents from servers over the World Wide Web. This is a huge global network that is used for education, entertainment, commerce, trade, information and many other purposes. HTML, although a text document, may link to other forms of media such as sound, image, streaming video. Web browsers process all these forms of media to present the end user with the resulting hypermedia that may include multiple forms of media. Moreover, web browser clients are supported on a range of devices, including desktop computers, personal digital assistants (PDAs), mobile phones, televisions sets and even Internet enabled Refrigerators.

User Model

Wahlster and Kobsa (1986) define the user model as the system's set of beliefs about the user, such as their goals, interests, knowledge, background and preferences. In this thesis we use the term *user model attribute* to refer to a single belief that is stored about the user in their model.

[1] History of the World Wide Web provided by http://www.w3.org/MarkUp/historical

[2] Source: http://www.orafaq.com

[3] Refer to http://www.w3c.org, accessed 6 January 2006.

One Modelling Conference' Reader's guide (Jameson et al., 1997) lists the following purposes of user modelling:

- Helping the user find information;
- Tailor information presentation to the user;
- Adapt an interface to the user;
- Choose suitable instructional exercises or interventions;
- Give the user feedback about their knowledge;
- Support collaboration;
- Predict the user's future behaviour.

This thesis focuses on the second point as this applies to the role of a user model in the personalisation of hypertext. A user model is an essential part of an adaptive hypertext system since it is the system's knowledge of the user and is the driving force determining exactly what is adapted and how.

The user modelling process determines the user model and its evolution over time. This can range from a very simple form, such as the user always setting the values of some flags in what is often called customisation, based on data elicited directly from the user in a questionnaire. At the more complex end of the spectrum, it may involve machine learning, based on information inferred or observed about the user (Schwab & Kobsa, 2002) or it may involve information and knowledge about other users such as stereotypic users (Rich, 1979) and deep knowledge of the domain. Jameson et al. (1997) provides a more exhaustive list of methods for constructing the user model and relevant literature.

Adaptive Hypertext and Hypermedia

Adaptive hypermedia is an alternative to the traditional "one-size-fits-all" approach in the development of hypermedia systems. Adaptive hypermedia systems build a model of the user's goals, preferences and knowledge of each individual user, and use this model throughout the interaction with the user, in order to adapt to the needs of that user (Brusilovsky, 1996; 2001).

Adaptive Hypermedia is a field of research that has grown in the past decade, evolving from a series of workshops to the International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (De Bra et al., 2004; 2002;[4]).

Brusilovsky divides Adaptive Hypermedia into two main categories: Adaptive Presentation and Adaptive Navigation support. He defines these terms as:

- Adaptive presentation - the adaptation of content (text, images, hyperlinks or any HTML element) of an hypertext document to a certain class of users. Brusilovsky's taxonomy

[4] Refer to Adaptive Hypertext and Hypermedia homepage online at http://wwwis.win.tue.nl/ah/

includes natural language adaptation, canned text adaptation (inserting/removing content fragments, altering fragments, stretchtext, sorting fragments, dimming fragments).

- Adaptive navigation - the adaptation of hyperlinks in a hypertext document to provide navigational support and prevent users from becoming lost in hyperspace. Brusilovsky's taxonomy includes several adaptive navigation techniques: direct guidance, sorting, hiding and annotation.

Whilst the two areas are certainly separate forms of adaptation, it is important to note that the adaptive presentation techniques allow, to a limited extent, navigation to be adapted: for example, to add or remove hyperlinks from the personalised document.

In this thesis, we use the term *Adaptive Hypertext* as in Brusilovsky's definition of Adaptive Hypermedia (Brusilovsky, 1996; 2001). Specifically, we mean the end result is an adapted, or personalised HTML document which can include text, images, sound and other forms of media that are supported by web browsers.

The term *personalisation* is used to describe an adaptation that has been performed by the system to match the user's specific interests, goals and other personal characteristics.

Scrutable Adaptive Hypertext

Kay (1999) defined *scrutability* in terms of the types of questions a user might be able to answer after scrutinising their user model to determine the:

- *values* of parts of the user model;
- *processes* that contributed to these values;
- *meaning* of the user modelling ontology.

The aim of this thesis is to extend this notion to consider how to make the *adaptivity* and *associated processes* open to the user and controllable.

The purpose of a *scrutable adaptive hypertext* is to ensure that the user can scrutinise the personalisation to understand and control it. Essentially, the user of an adaptive system should be able ask the system *How was this material personalised to me?* and to be able to see just how their user model affected what they saw in the hypertext document. With an understanding of the personalisation process and the ability to control it, the user has the foundations of control over personalisation as they can steer the personalisation as they choose.

1.2 Aims and Contributions

This thesis is an exploration of how to provide effective support for user scrutiny of personalisation of an adaptive hypertext. It extends on the work of Kay (1999) with scrutable user models, by adding support for scrutinising the process of personalisation in addition to the user model. This involves two quite different classes of challenge. On the one hand, it requires the design of an architecture for an adaptive hypertext system that makes it possible for the user to scrutinise the adaptation processes. This poses considerable technical challenges. The other, quite different class of challenge is the creation of a suitable interface that supports users in scrutinising the personalisation.

We now describe the main contributions of this thesis.

Creation of an architecture for an adaptive hypertext system that makes it possible for the user to scrutinise the adaptation processes

This thesis utilises a standard adaptive hypertext technique and builds on it by adding support for the scrutinisation of the personalisation it provides. The personalisation mechanism is limited to adaptive addition and removal of content fragments based on rules embedded in the hypertext document requested and the user model of the user to whom the page is personalised. This is simple but also representative of typical adaptive hypertext systems (Brusilovsky, 2001).

The extension of this well known adaptive hypertext technique to support scrutiny of the personalisation is a contribution of this thesis. This thesis contributes SASY and ATML, which collectively form a framework for the generation of adaptive hypertext documents in such a way that the personalisation may be scrutinised by the user.

Creation of a suitable user interface that supports users in scrutinising the personalisation of an adaptive hypertext

This thesis makes a contribution in its novel user interface through which the user may scrutinise personalisation. The novelty is that it enables the user to see the implications of their user model in terms of the adaptations that it effectively controls.

In this thesis, we use the term *scrutinisation tools* to refer to the components of SASY that allow the user to scrutinise a personalised hypertext in the following ways:

- understand exactly what aspects of the hypertext were personalised. That is, what parts of the content were added to and/or removed from the user's personalised view of the page;
- what aspects of the user model caused each personalisation;
- view evidence of how user model attributes were set, whether set directly by the user or inferred by the system;

- understand and visualise the implications of changing their user model, in terms of how it affects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes;
- change aspects of their user model to control the personalisation.

This thesis explores ways to support the user to scrutinise simple and more complex forms of content personalisation.

We define s*imple* content personalisation as:

- Personalisation based on the value of a single user model attribute.
- Personalisation based on user model attributes set explicitly by the user, such as a user preference.

We define *complex* content personalisation as:

- Personalisation based on the result of expressions involving one or more user model attributes.
- Personalisation based on user model attributes that have been inferred about the user.
- Personalisation that changes the way content is presented to the user over time based on a changing and evolving user model.

This distinction is particularly important for the user interface issues. Support for the process of scrutinisation of both simple and complex content personalisation is supported by the ATML framework as well as the SASY scrutinisation tools. The boundary and scope of the complexity supported and how it relates to other forms of adaptive hypertext is discussed in the next section.

Evaluation of the Scrutinisation Tools

This thesis contributes a detailed evaluation of the scrutinisation tools across two different domains to reduce the effect of a particular domain on the results. In addition, the evaluations demonstrate the application of scrutable adaptive hypertext to different types of adaptive systems: an adaptive teaching system, an adaptive recommendation system. The evaluations explore scrutinisation of simple and complex forms of personalisation.

- The first evaluation reports a field study, where students used SASY to learn about UNIX Security in an authentic environment. In this study SASY was used as an adaptive hypertext teaching system. The evaluation presents insights into ways users scrutinise in a real world application, where their main goal for using the system is to learn about UNIX Security. It analyses the most-used and least-used forms of scrutinisation and the typical usage patterns.

- The second evaluation reports a laboratory based experiment where we asked users to complete tasks around a Personalised TV Guide. Here SASY was utilised as an adaptive recommender system that recommended a personalised television program viewing schedule to users based on their specific interests. The experiment evaluates the usability and effectiveness of the scrutinisation tools. The report presents qualitative data about the use of the scrutinisation tools and user feedback on the notion of scrutinisation of adaptive hypertext in general. It also provides insights into user attitudes about personalisation based on inferences and bold assumptions about the user.

Guidelines for providing support for the scrutinisation of an adaptive hypertext through the exploration of several forms of scrutinisation tools

There are several challenges in creating scrutinisation tools:

- It is not obvious how to present and explain personalisation to the user in way they can easily understand.
- The tools must be easy to find and access when required, even though users may not use them very often.
- It must support different types of users. Some users may scrutinise often, others might scrutinise rarely, for example, when the personalisation produces unexpected results.
- The tools must be easy to use without any training, since people will probably use them on a needs basis.

This thesis contributes a set guidelines to overcome these challenges, based on outcomes of the development and evaluation of several forms of scrutinisation tools explored by this thesis.

Overview of the Scrutinisation tools explored

We began with the expectation that the main challenges in supporting scrutably adaptive hypertext would be technical. Accordingly, we began with the simplest form of adaptive hypertext, single attribute controls presentation of text fragments, using Tutor 1. Our plan was to explore how to support scrutable adaptive hypertext for this simple case and then, once this was quickly achieved, we would move onto the more challenging case where the user model and the processes of adaptation were more complex.

At this time, we presumed that the user interface design should involve a very subtle addition to the adaptive hypertext page, a simple link to click when the user wanted to scrutinise the adaptation. Our reasoning was that most of the time, most users would be primarily engaged in doing the activity that was supported by the adaptive hypertext. On the occasion that the user decided they wanted to scrutinise, we expected they would look for our subtle link and follow it through.

In hindsight, we underestimated the power of user expectations. People have become accustomed to computers that adapt to their behaviour to the user. However, people are NOT used to being able to understand that adaptation: they normally have to accept it. For example, a study of a personalised web site MyYahoo! (Manber et al., 2000) claimed most users stay with the default presentation and never customise.

Essentially, our conclusion at this stage was that people report liking the idea of control over personalisation but they are not ready for it.

After user trials with our first series of scrutable adaptive hypertexts (Tutor1 1, Tutor 2 and Tutor 3), we realised that a major challenge was in the user interface and that we needed to find ways to deal with people's existing mental models which include the belief that computers adapt but that one cannot understand or scrutinise the way this works.

The main breakthrough at this stage was to explore making the personalisation and its control much more obvious in the user interface of Cell-Tutor (developed by Serena Potts and Judy Kay, 2005a, 2005b): all user model attributes that figure in the personalisation were presented at the right of the interface and nearby there are links to enable other tools for scrutiny.

This interface was strikingly more successful that previous Tutor series interfaces from the point of view of supporting scrutability. We extrapolated this idea of a highly visible scrutability interface and then explored more complex personalisation in SASY. This is the subject of the thesis.

Table 1.1 describes the key differences between the systems created and evaluated as part of this thesis. SASY 4 and ATML 2 are the last versions of a series earlier systems developed and studied as part of this research:

- Tutor 1 (Czarkowski, 1998; Czarkowski & Kay, 2000);
- Tutor 2 (Czarkowski & Kay, 2001, 2002);
- Tutor 3 (Czarkowski & Kay, 2003a, 2003b & 2006);
- Cell-Tutor (Czarkowski, Kay & Potts, 2005a, 2005b);
- SASY 1 (Czarkowski, 2005c);
- SASY 2;
- SASY 3.

This thesis focuses on SASY 4 and ATML 2 for the sake of brevity but presents the lessons learnt from the previous systems in terms of providing support for the process of scrutinising personalisation of an adaptive hypertext.

Table 1.1 - Summary of key differences between systems developed, as part of this thesis, to explore support for the scrutinisation of personalisation of an adaptive hypertext.

System	User Model Complexity [5]	Adaptive Hypertext Complexity	Visibility	Evaluation		Comment
				Quant-itative	Qual-itative	
Tutor 1 (ATML 1)	Simple	Simple	Low	✔		Initial work with scrutable adaptive hypertext.
Tutor 2 (ATML 1)	Simple	Simple	Low		✔	Improved user interface to scrutinisation tools.
Tutor 3 (ATML 1)	Simple	Simple	Low		✔	Re-designed user interface to scrutinisation tools.
Cell-Tutor	Simple	Simple	Moderate		✔	Increased visibility of user model.
SASY 1 (ATML 2)	Complex	Complex	High	✔		Increased complexity of user model and personalisation. Increased visibility of scrutinisation tools.
SASY 2 (ATML 2)	Complex	Complex	High		✔	Slightly refined user interface to scrutinisation tools.
SASY 3 (ATML 2)	Complex	Complex	High		✔	Slightly refined user interface to scrutinisation tools.
SASY 4 (ATML 2)	Complex	Complex	V. High		✔	Increased visibility of scrutinisation tools.

1.3 Scope and Limitations

This section discusses the scope of this thesis and sets to boundary of our work. We discuss the class of adaptive hypertext systems for which we explore scrutability; the complexity of the adaptive response that is made scrutable; adding scrutability to other adaptive hypertext frameworks and issues with authoring scrutable adaptive hypertexts.

Adaptive Hypertext

As mentioned above, this thesis focuses on supporting scrutability in a certain class of adaptive systems: adaptive hypermedia systems. There are a couple of reasons for this. Firstly, there is an active and growing research community working with this technology (De Bra et al., 2004; 2002). Secondly, the technology is widely used, ranging from online retailers, information search and retrieval to education. For a comprehensive list of example applications refer to (Brusilovsky, 2001).

[5] Note that we use the term complex according to our definition earlier in this chapter, which is not as complex compared to some other forms of adaptive system, as discussed in the Limitations section.

The mechanism used to generate personalised hypertext is not the focus of this thesis. Rather standard adaptive hypertext techniques are employed as a basis for exploring scrutability of the personalisation. This thesis explores several forms of adaptive hypertext but concentrates on the scrutability of adaptive presentation, where the personalisation involves inserting/removing content fragments based on logical expressions involving values of user model attributes. While this is a simple approach, it makes a logical starting point for exploring ways to support the scrutinisation of an adaptive hypertext. The simple approach is nonetheless an effective means of personalisation, which makes it appealing and more likely to be used widely, compared to a complex approach with the same end result. Furthermore, it is similar to the mechanisms employed by several commercial application servers to provide personalisation services. For example ATG 7[6] and BEA WebLogic Personalisation Server 3.5[7] use rules defined by Business Analysts as the basis for personalisation of content. In fact, Höök (2000) points out that the "very few intelligent user interfaces that have succeeded commercially have done either very simple adaptations based on simple knowledge of the user, or created its adaptations based on what other users do rather than some kind any complex inferred model of user."

Complexity of Adaptive Response

We have stated this thesis explores *complex content personalisation*. However, it is important to note there are much more complicated approaches used to generate a personalised response: for example, personalised responses generated by a mechanism using Classification Learning, Collaborative Filtering, Bayesian Networks, Natural Language Generation (Jameson, 2003b). Figure 1.2 shows the relationship between the systems explored in this thesis and other forms of adaptive systems. Tutor 1 / ATML 1 used a simple user model and simple rules for personalising content. SASY 4 / ATML 2 extended this initial work to more complex forms of both. One might imagine that an adaptive teaching system might have an advanced adaptation approach, based on pedagogical planning and reasoning. A recommendation system might have a more complicated user modelling approach. This thesis does not claim to cater for the scrutinisation of yet more complex adaptation such as natural Language Generation. However, what is called *complex* personalisation in this thesis is thesis is both useful and powerful, and thus makes a good starting point to explore ways to support the scrutiny of personalisation. While it is possible that research presented in this thesis is applicable to more complicated methods for personalisation, it may not necessarily be the case.

[6] For ATG 7 product documentation refer to https://www.atg.com/repositories/ContentCatalogRepository_en/manuals/ATG7.1/business/index.htm accessed 27 November 2005.

[7] For BEA WebLogic product documentation refer to http://e-docs.bea.com/wlcs/docs35/index.htm accessed 27 November 2005.

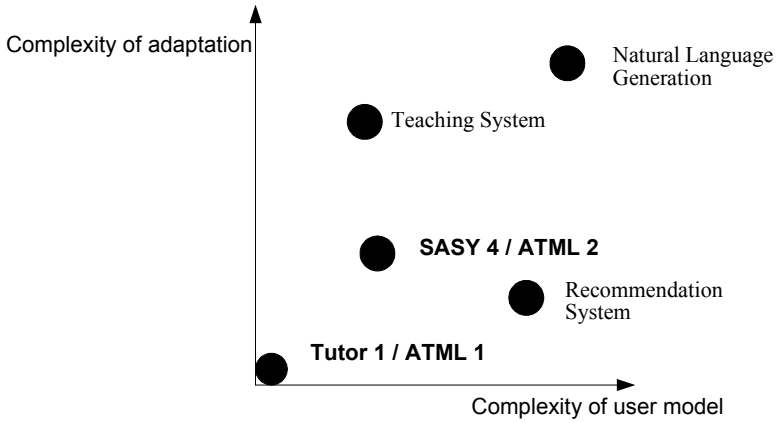


Figure 1.2 - Graph showing relationship of complexity of the user model and adaptation considered by this thesis to some other forms of adaptive systems.

Increasing the complexity of the personalisation mechanism reduces the predictability and transparency of the personalisation to the user (Jameson, 2003) and undoubtedly increases the complexity of explaining the personalisation to the end user. For example, if a page fragment is included as a result of a complex rule involving 100 user model attributes, it is not obvious how this personalisation would be explained to the user and how they might control it. There is a trade-off between the complexity of the personalisation and its scrutability. However, if a personalisation rule is difficult for a human to explain, it would also be difficult to develop, test and ensure its correctness. For this reason, we can expect that handcrafted personalisation will tend to be at the simple end of the spectrum. Moreover, even automatically generated adaptive hypertext may perform useful personalisation that is quite simple. So the limits imposed upon the scope of this thesis still accommodate a range of useful and interesting applications. Importantly, we need to gain better understanding of the simpler personalisation first: it seems unwise to tackle the complex without that.

To support more complex personalisation, this thesis follows Höök's "black box in a glass box" approach (2000; Höök et al. 1996): arguing what is important for scrutinisation is the ability to provide the end user with an understanding of the overall result of the personalisation and how it might be changed to achieve a desired affect, rather than an in-depth understanding of the algorithm. To support this, SASY/ATML allows the content author to provide explanations for complex personalisation rules. For simple rules, explanations are generated by the system.

Creation of a Scrutable Adaptive Hypertext framework

This thesis argues that scrutability support should be considered by adaptive hypertext systems right from the system design phase. In this approach, scrutability is at the core of the system and content created by authors of personalised content. For this reason, this thesis uses its own adaptive hypertext system rather than using on existing one. Exploring how to add scrutability to an existing adaptive hypertext system presents a different set of challenges and is beyond the scope of this thesis.

Authoring of Scrutable Adaptive Hypertext

The focus of this thesis is on the end user to whom the hypertext is personalised. As such, the process and issues involved in authoring the adaptive hypertext is beyond the scope of this thesis. However, scrutability of the adapted hypertext could benefit the author as well as the end user, as the author is responsible for ensuring the correctness and quality of the hypertext. Scrutability provides a mechanism for checking the adaptive hypertext does in fact adapt as intended. The issue of authoring is important, but is beyond the scope of this thesis.

1.4 Examples of a Scrutable Adaptive Hypertext

The section describes concrete examples of how a user might scrutinise an adaptive hypertext as proposed by this thesis. Consider Scenario 2 below, which is also the basis of Evaluation 1 in Chapter 6.

Theresa uses an adaptive hypertext teaching system to learn about UNIX File System Security. She is already quite familiar with UNIX, but is required to use this resource as part of her homework assignment. On accessing the system she completes a quiz and the system personalises it teaching of UNIX File System Security to her.

Scenario 2 – Scrutinising personalisation of an adaptive hypertext teaching system

This represents a typical scenario in a personalised hypertext system. The system stores a model of the student, including their knowledge, and presents material to the student based on pedagogical rules about the teaching domain. This thesis is concerned with how the student might go about scrutinising the personalised material presented to them.

The student might ask the following questions:

1. Why does it think I do not know about the Chmod command?
2. What else does it think I know and don't know?
3. Why does it show me the Chmod command when I already know this?
4. What did it not show me?
5. Why didn't it show me how to use the numeric form of the Chmod command?
6. Why did it explain the Chmod command the way it did?

The first two questions can be answered by scrutinising the user model itself, for example using the approach described by Kay (1999). The user can access and scrutinise the user model to find out what the system thinks it knows about them: what it thinks the student knows, what it thinks the student does not know; and examine evidence for these beliefs.

This thesis builds on the work of Kay by adding support for the user to scrutinise the process of personalisation rather than just the user model. The user can gain partial information about the process from the user model, but they also need information about how the user model was used to complete the picture. For example, to answer the third question, the user could scrutinise their user model and investigate why the system thinks they do not know the Chmod command. Another approach is to scrutinise the personalised content directly to determine why the paragraph about Chmod was presented by the system. In other words, the meaning of the user model may well be best explained in terms of the way it is used for personalisation.

Furthermore, there are some forms of scrutinisation that are external to the user model, for example, the fourth question. To answer this, the user needs to be able to query the personalisation system to find out what content was removed from their view. The user might want even more information. For example, in the fifth question, having found out a particular form of the Chmod command was not shown to them, they might want to know why? Likewise, as in the sixth question, they might want to know why information that was shown to them was presented in the way it was. Perhaps it was too verbose, or not explained enough or did not provide the right level of background? The user might want to see what alternate forms of explanation are available. Consider the case where the user has specified a preference to be shown only simple quiz questions rather than more complex ones. There may be a discrepancy between what the user regards as simple as opposed to what the system classifies as simple. The ability to scrutinise the personalised content allows the user to potentially fine tune the personalisation to better suit their needs. This increases the accuracy of the user model and should improve the accuracy of future personalised responses.

As another example, consider the following scenario:

Halina logs on to an online, personalised recipe repository to look up the recipe for tuna rissoles. The recipe is personalised to her based on the dietary requirements and preferences she has specified in her profile.

Scenario 3 – Scrutinising personalisation of an adaptive hypertext recommender system

Having seen the recipe, Halina might wonder:

1. Why does the recipe include margarine based on canola oil?

2. What did it personalise in this recipe to suit my low cholesterol dietary requirements?

3. What ingredients would the recipe include if the system thought I was not concerned about cholesterol?

4. How can I make it show butter instead of margarine?

It is not entirely obvious why the system included a canola oil based margarine in the recipe. Scrutinising the user model might provide clues but no definitive answer. In this case, the user needs to be able to query the personalisation system to find out *why* it was included. Scrutinisation might reveal that this was because Halina has low cholesterol dietary requirements. She might wonder what else the recipe included to suit this requirement.

The third and fourth questions relate to the issue of control over the personalisation. If Halina is preparing to host a dinner party, she might wish to cook with full fat ingredients for her friends who appreciate taste over cholesterol concerns. She could directly change her user model to select full fat ingredients over low cholesterol substitutes, but this might affect recipes she is shown later on. She might wish to first understand the impact of changing her user model as there may be adverse effects she is not aware of. Or perhaps she might just wish to see an alternate version of the recipe on this occasion only, but retain her dietary requirements for future use.

The fourth question relates to user control of the personalisation, given the user has made a decision to change it. The user could directly access the user model, but again is may not be clear what it is they need to change. Scrutinising the personalisation, as in Questions 1 and 3, might reveal butter as an alternative ingredient. It should also explain what is necessary to change the personalisation to include butter rather than margarine in the recipe. This would provide guidance to the user and tell them how to change their user model to achieve the required result.

1.5 Motivation

This section discusses the motivation for this research.

Usability issues: Predictability, Transparency, Control and Unobtrusiveness

The emergence of intelligent and adaptive systems started the debate that these systems, where not carefully designed, violate many of the good usability principles developed for direct-manipulation systems including control, transparency, predictability, privacy and trust (Höök, 2000). A debate on direct manipulation interfaces vs. interfaces driven by software agents highlighted understanding and control as key usability issues with adaptive systems (Maes, Schneiderman, 1997). An agreed outcome of that discussion was that users need to understand the adaptive operation of an adaptive system, to some degree, in order to trust it to perform tasks on their behalf. In addition, users must feel as though they have ultimate control over the system if and when they choose to exercise it. Essentially, users might want to know what the system knows about them and how it uses this information. Jameson (2003) argues that if a system explains its adaptive actions to the user and allows them to inspect their user model, this compensates for the reduced transparency and predictability of the adaptive system. Scrutinisation of personalisation is just that. It allows the user to understand why personalisation occurred in a manner they are able to comprehend.

This thesis explores another dimension of predictability. It considers how a user might visualise the implications of changing their user model. This allows users to not only predict but actually see the effects of changing their profile before they make any changes.

The issue of control has further implications. On one hand, from a privacy perspective, the user should have access to personal data that a system stores about them: the ability to change it, the ability to change how it is used. On the other hand, particularly in an adaptive teaching system, there are potential dangers if learners are given unrestricted access to edit their user model (Kay, 2001; Tanimoto, 2005). The user could quite easily undo all the hard work the adaptive system has done to form its assessment of the user. Such is the argument that is put forward (Tsandilas & Schraefel, 2004) about a Microsoft intelligent agent, that supports direct manipulation by allowing users to configure its intelligent behaviour (Horvitz, 1999). Tsandilas & Schraefel argue that since the underlying adaptation model is not transparent to the user, as a result, the system's behaviour may appear inconsistent and unpredictable. One remedy for this is to restrict access to parts the user model. The approach used in Mr. Collins (Bull, Pain, 1995), is to force the user to negotiate changes to their user model. Mr. Collins allows the user to inspect their user model to see beliefs the system holds about them, offer their own self assessment and then negotiate with the system to change its beliefs about the user. The user has to correctly answer test questions to convince the system. Whilst it protects the validity of the user model, it limits the extent to which the user can control the system.

This thesis follows a different approach, supported by (Tsandilas & Schraefel, 2004). It explores how to empower the user with enough information about their user model and how it affects personalisation such that the user is aware, to some degree, of the impact of changing their model. Tanimoto (2005) argues one of most challenging current challenges is determining how much support is provided in explaining and interpreting learner models to the learner. Bull and Kay (2005) argue a learner needs to understand the effect of the model on personalisation. By involving the user in the personalisation process and allowing them to experiment with what-if scenarios to understand the implications of changing their user model: the risk they might make user model changes that corrupt or adversely affect the system's function may be reduced. Basically, scrutinisation helps the user to make an informed decision when changing their user model.

Kay (2001) points out users might not always be interested is scrutinisation. In fact, they may go for long periods without scrutinising. In a learning context, they might only scrutinise at the end of their use of an adaptive teaching system as a means of reflecting about their learning. They might stick with the default adaptation at all times, trusting the system to perform its expert function. Even so, users would take comfort in having some level of transparency in the system, knowing they are in control of the personalisation and able to change it or turn it off if necessary. However, Alpert et al. (2003) conducted an empirical evaluation that investigated user attitudes toward user-adaptive sites. In their study, users expressed a strong desire to have full and explicit control of data and interaction. Users wanted to be able to make sense of the adaptive web site's behaviour by understanding the rationale for displaying particular content. This is precisely the goal of this thesis. It is also consistent with the observation that many people want to control the use of personal data (Ackerman et al., 1999), especially where that data can be linked to their identity. There are also recent studies that report users also want to control the personalisation itself (Bontcheva et al., 2005; Jameson, 2005; Peng et al., 2005; Cheverst et al., 2005).

Scrutability also has implications for obtrusiveness. The explanation of personalisation to users involves a trade-off between the comprehensibility of the explanations their obtrusiveness. The visibility of the scrutinisation tools is a key issue explored in this thesis. For example, if users do not scrutinise often, the scrutinisation tools should not interfere with their main task for using the adaptive system in the first place, which might be to learn or find information. However, when they need to understand what caused a particular personalisation and take control, they should easily be able to do so. It is not clear where how obtrusive scrutinisation support should be and this thesis explores how to find the right balance.

Privacy and Legislation

A review of consumer polls regarding privacy issues (Kobsa, 2002) reported users are very concerned about the use of their personal data. If asked for personal details, some users would fake registration details or leave a website altogether. One way around this is to build the user model by unobtrusive

means such as inferences made about the user (Schwab & Kobsa, 2002). For example, Amazon (www.amazon.com) bases personalised product recommendations based on the consumer's previous purchases and rated products. User models that are based entirely on inferences may suffer from inaccuracy, as described in the next section.

Encouragingly, some surveys on personalisation (Kobsa, 2002; Personalisation Survey, 2000) claimed users were willing to give out personal information for something valuable in return. According to Kobsa, providing means to allow users to inspect, block, rectify, erase both the data that they themselves provided and specifically the assumptions that the system inferred about them, would be likely to increase users' trust in the application.

It appears that internationally, legislation embodies the view that this kind of access to personal data is desirable: for example, the European Directive for Data (European Parliament Privacy Directive, 1995; Kobsa, 2002). Its Article 12, "Right of Access", states users of systems that store personal information about them must have:

- (from part a) communication in an intelligible form of the data undergoing processing and of any available information as to their source;
- (from part a) knowledge of the logic involved in any automatic processing of data concerning the user;
- (from part b) a means for rectification, erasure or blocking of data the processing of which does not comply with the provisions of the Directive, in particular because of the incomplete or inaccurate nature of the data.

This thesis argues that to fully support the spirit of this legislation, the adaptive system should go beyond making personal data available to the user, to involve the user in the personalisation *process* and it should empower them to understand and potentially alter it to better suit their needs. This requires the system to support the user not only to access, understand and modify data that is stored about them, but also, to enquire, understand and control *how* that data is used and processed to produce the personalised response.

Benefits gained by correcting misconceptions and errors

Adaptive systems are not perfect. They may hold misconceptions about the user, misconceptions about domain, errors in the adaptive content or even program bugs. Scrutinisation helps the user to understand data stored in their user model and how it is used. This allows the user to validate aspects of their user model and correct misconceptions and errors the system has about them.

Adaptive systems that rely on unobtrusive methods for developing user model are particularly susceptible to this. Kay (2001) argues that making the user model scrutable has the added benefit that correctness of the model can be assessed by the user; this in turn helps address one of the objections to

user-adapted systems on the grounds that they may infer too much from too little information (Henderson, 1995).

In contrast, a paper about MyYahoo! (Manber et al., 2000) claimed most users stay with the default presentation and never customise. That is, users were not prepared to spend time inspecting and populating their user model. This was claimed to be due to a combination of the default page being good enough to not require personalisation and the effort involved in learning how to customise the site. However, it is also possible that users may have been concerned about altering things without knowing the impact of their changes. Indeed, for some non critical applications, such as a personalised online newspaper, this might be the case. The internet provides a huge repository of information. If users can not find what they are looking for on one web site, it is easy to try at another site. Users may not want to understand or control the personalisation, or may trust the system to perform its personalisation function. For example, in the case of a teaching system, learners may prefer the teacher (the system) takes control (Kay, 2001). Nonetheless, it is possible the teacher (whether human or machine) might make an incorrect assumption about the student. For example, it might start teaching them something they already know very well. At this point the student might at least want to tell the teacher they wish to skip this topic because they already know it.

If the application is one that the user has no choice but to use, for example, it is part of their job, or the user sees value in fine tuning the personalisation, they might be motivated to scrutinise and correct the user model and personalisation. At the extreme end, users might proactively wish to volunteer information about themselves in return for better service, particularly if they frequently use or depend on the service. One example of this occurs in Amazon (www.amazon.com) where consumers can tell the website they do not like certain products and that the personalisation should not use a particular purchase as evidence they like the product. Amazon also enables consumers to rate products, adding valuable information to the user model that, in turn, helps to make the personalisation more accurate. As another example, Manber et al. (2000) also reported there were some “power” users who had spent a great deal of effort in customising the site, building customised pages bigger than 500KB with stock portfolios of more than 200 stocks. Forrester research (Johnson, 2004) claims that, as the online retail industry has reached profitability and sales have begun to stabilise, retailers are investing in online selling innovations: retailers now use more explicit customer information to deliver personalized shopping experiences by giving consumers the ability to actively pull more relevant information and tell retailers when they are seeing products and content that they do not like.

Also, in the case of critical systems, there may be more motivation for users to scrutinise and have traceability and control over the system’s adaptive response. These are systems that make critical decisions or perform critical tasks on behalf of their users. Consider, for example, a system that invests its user’s savings, based on their investor profile and risk adversity.

From a different perspective, another reason for scrutability relates to the possibility of errors in the adaptation. The more complicated the hypertext, the less efficient the authoring (Calvi & Cristea, 2002) and the more opportunities for authoring errors. One approach to this problem is to improve authoring tools. This is undoubtedly an important field of research. For example, tools have been developed for AHA! (De Bra et al. 2003). While AHA! has a similar architecture and adaptivity features to that of SASY, its authoring tools include a specialised editor to define concepts and adaptation rules and a Graph editor which allows the author to visually link concept nodes (De Bra et al. 2003). Such authoring tools will help reduce the difficulty in authoring adaptive hypertext and also reduce the number of errors. However, it is not possible to completely eliminate mistakes. In fact, a scrutable interface could be used as a debugging tool by the adaptive hypertext author as it should clearly explain how a page was adapted to the user.

A closely related problem can occur where there is mismatch between the user's understanding and the author's intention. For example, an adaptive maths course may offer the user the choice of an easy or more challenging version of the course. A user might request the easy version, only to find it is too easy and that they would have actually preferred more challenging material.

Curiosity, exploration and adventure

Another reason to provide users with the ability to scrutinise personalisation is to support their sense of curiosity and exploration. The user can exploit the scrutinisation to allow them to explore alternatives. For example, in a personalised holiday recommender system, the user might wonder what holiday packages the system did not show them and what would have been included if the user were able to tell the system they really like the beach?

As another example, consider a teaching system that personalises based on the user's learning style. The system might suggest an appropriate teaching strategy based on a psychological analysis of the learner, or perhaps the learner has chosen a particular strategy. The learner might still be curious to know how the material might be presented if they chose a different strategy?

From another perspective, the user might simply wish to understand how the system personalises information to them. Knowing this might satisfy their curiosity or simply provide comfort and perhaps develop trust by giving them a sense of understanding and control.

Reflection about learning

There is a substantial and growing body of work that makes the user model available to the user (Bull, Brna, Pain, 1995). In particular, in educational contexts, there has been considerable interest in the potential of supporting reflective learning by creating suitable interfaces to the user model as reported, for example, in a series of workshops (Kay, J., Lum, A. and Zapata-Rivera, 2005; Bull, Brna,

Dimitrova, 2003; Morales, Pain, Bull and Kay, 1999). Even in systems where the work is not explicitly concerned with supporting student reflection, there are many cases where the learner is given some insight into the system's model of the learner. For example, Corbett and Anderson (1995) describe a teaching system that provides a set of 'skillometers' which show a summary of the learner model in terms of the degree to which a set of skills has been learnt. Of course, displaying the user model provides the user with access to some understanding of just one of the core elements of the personalisation. Scrutinising the personalisation might encourage the user to reflect on their user model in general.

1.6 Summary and Thesis Overview

This thesis argues the importance of taking adaptive systems beyond making personal data available to the user, to involve the user in the personalisation *process* and empowering them to understand and potentially alter it to better suit their needs. This requires the system supporting the user to, not only access, understand and modify data that is stored about them, but also, to enquire, understand and control *how* that data is used and *processed* to *produce* the personalised response.

We have explained our motivation for a scrutable adaptive hypertext:

- compensates for the lack of predictability, transparency and controllability in adaptive hypertext systems;
- provides a mechanism to support the spirit of legislation that requires users to not only have access to personal data stored about them, but also the support to understand how that data is used by the system;
- allows users to see the implications of changes to their user model;
- allows users to experiment with "what-if" scenarios exploring alternate forms of personalisation;
- allows users to validate aspects of their user model and correct misconceptions and errors the system has about them – this has the benefit of improving the accuracy of the user model and, in turn, improving the personalisation;
- helps users satisfy their curiosity and exploration of how the personalisation works;
- may encourage reflection of the personalisation and user model in a learning context;
- the scrutiny of the personalisation may be the most accurate way for a user to understand the pragmatic meaning of the user model.

This thesis explores ways to support the process of a user scrutinising a hypertext that has been personalised to them. It extends the work on scrutable user models, by adding support for scrutinising the process of personalisation in addition to the user model. The contributions of this work are:

- The creation of a framework for a reusable, domain-independent, adaptive hypertext system that makes it possible for the user to scrutinise the adaptation processes.

- The creation of a usable and effective set of graphical scrutinisation tools that allow the user to see the implications of their user model in terms of the adaptations that it effectively controls.

- Evaluation of the final scrutinisation tools, across different domains, also demonstrating the application of scrutable adaptive hypertext to different types of adaptive systems.

- A set of guidelines for the development of scrutinisation support based on outcomes of the evaluation of each form of the scrutinisation tools developed as part of this research.

The structure of the remainder of this thesis is:

- Chapter 2 presents related work;

- Chapter 3 introduces the framework for a scrutable adaptive hypertext. It describes the requirements for a scrutable adaptive hypertext, the SASY / ATML architecture and scrutinisation tools in SASY 4;

- Chapter 4 describes ATML 2, the language used by authors to create adaptive hypertext that the user can scrutinise;

- Chapter 5 discusses the process of exploration of support for scrutinisation. It presents the evolution of the systems created, evaluated and used to inform this research: Tutor 1, Tutor 2, Tutor 3, Cell-Tutor, SASY 1;

- Chapter 6 reports the evaluation of the scrutinisation tools, and explains the refinements made to the scrutinisation tools through a qualitative evaluation to create SASY 4;

- Chapter 7 presents the guidelines for providing support for the process of scrutinising an adaptive hypertext and conclusions.

Chapter 2 Background

This thesis is an exploration of how to provide effective support for user scrutiny of personalisation of an adaptive hypertext. The primary goal of this chapter is to establish the context of our work.

This research relates to two separate, but related fields: Adaptive Hypertext and Hypermedia (AH) and Open User Modelling. This chapter presents a brief background for AH to provide context for the class of adaptation that we aim to make scrutable through this work. We then discuss related research in to Open and Scrutable User Models and survey the literature to review the current extent of scrutinisation. We discuss the literature in relation to our framework for scrutable adaptive hypertext: SASY (Scrutable Adaptive hypertext System) and ATML (Adaptive Tutorial Mark-up Language).

2.1 Adaptive Hypertext

Adaptive Hypertext has it origins in User Modelling, in particular Adaptive Systems, and Hypertext. Hypermedia was first proposed by Vannevar Bush in 1945. Bush provided his vision of browsing the Web of linked information, which included the ability for readers to add their own information to add to the growing Web. According to Brusilovsky (2002), "pre-Web generation of adaptive hypermedia systems explored mainly adaptive presentation and adaptive navigation support and concentrated on modelling user knowledge and goals". There was particular focus on using hypermedia for education, for example (Nunes et al., 1996; Hohl, et al., 1995), information access and help (see Brusilovsky 1996 for a list of works). This research made important progress until the mid 1990's, when the World Wide Web (WWW) became a popular mechanism for the delivery of a new form of hypertext documents, HTML (HyperText Mark-up Language). The WWW added flexibility as author's could easily create and publish HTML documents to the WWW, that could reference other forms of media (collectively, termed hypermedia).

Soon after the emergence of the WWW, the literature reports research that explored how to apply the methods that had been developed for User Modelling and Adaptive Systems to hypermedia (Brusilovsky, 1996). The initial drivers were to develop ways to tailor the hypermedia to the individual user to best suit their needs, and to support the user as they navigated through hyperspace, so that they did not get lost, that is, become overloaded with information or lose track of their initial goal having followed unrelated hyperlinks. The research lead to the establishment of a new field: Adaptive Hypertext and Hypermedia (AH), with a series or workshops and international conferences: workshop at UM'94 (Fourth International Conference on User Modelling)[8], AH2002 (De Bra et al., 2002), AH2004 (De Bra et al., 2004).

[8] Refer to http://wwwis.win.tue.nl/ah94/, accessed 24th February 2006.

At a simplified level, Adaptive Systems use information about each individual user, called a *user model*, to adapt their behaviour to that user. The user model stores personal characteristics about users such as their the goals, interests, background and knowledge. According to Kobsa (2001), research into user modelling can be traced back to the late 1980's with the works of Allen, Cohen and Perrault (Perrault et al., 1978; Cohen and Perrault, 1979; Allen, 1979) and Rich (Rich, 1979a, 1979b). Research in to Adaptive Hypertext and Hypermedia (AH) is focused at applying known methods and techniques for user modelling and adaptive systems, and developing new ones, for the new medium of hypermedia. This includes constructing, maintaining user models, exploiting information about users to provide an adaptive response, developing ways to adapt the content and presentation of hypermedia to the individual user and support their navigation through it.

We will not focus on the specific methods, techniques and issues for user modelling and adaptive systems as they are not the focus of this work. However, it is useful to describe AH in terms of the broader context adaptive systems. We do so here following a framework used by Jameson (Jameson et al., 1997; Jameson, 2003;2003b), that describes the key dimensions of user-adapted interaction: Purpose of Adaptation, Data used for adaptation, Methods for constructing the user model, Methods for exploiting the user model.

Purpose of Adaptation

As described by (Jameson, 2003) Adaptive Systems have been developed for a range of purposes including, but not limited to, the following:

- Tailor information presentation to the user;
- Adapting an interface to the user;
- Giving advice about system use, for example, just in time help;
- Helping users find information;
- Making personalised recommendations to the user, for example recommending products, music, movies, a navigation path through a physical space such as a shopping centre;
- Supporting learning;
- Supporting collaboration;
- Taking over parts of routine tasks;
- Controlling a natural language dialog.

Brusilovsky's surveys of AH (1996, 2001) reported the main purposes of AH systems as: educational hypermedia, on-line information systems, on-line help systems, information retrieval hypermedia, institutional hypermedia, and systems for managing personalized views in information spaces. In the 2001 survey, Educational hypermedia and on-line information systems were established leaders (accounting for two thirds of the research) with Information retrieval hypermedia close behind.

Data used for adaptation

This dimension describes what sort of information is stored and used by the system to produce an adaptive response. We extend Jameson's description of this dimension to include not only information about the user, but also information about the environment and usage data. Naturally, the type and depth of information modelled is driven by the purpose of the adaptation. For example, an AH teaching system will typically model domain knowledge, whereas a movie recommendation system might model user interests instead. Brusilovsky's surveys of AH (1996, 2001) describe the following types of data being modelled by AH systems and using this data to provide a personalised response to the user: user's goals, tasks, domain knowledge, background, hyperspace experience (i.e. familiarity of the hyperspace structure), preferences, interests and individual traits such personality type (e.g. introvert/extravert), cognitive factors, learning styles, environment factors such as the user location (e.g. direction of sight, movement through a physical space such as a shopping centre or museum) and the user platform (e.g. hardware, software, network bandwidth), usage data, for example, purchase history (Kobsa & Josef, 2000). Environmental attributes have become more important recently with the increasing commercial uptake of mobile, hand-held devices that are connected to the WWW, and emerging research in to ubiquitous computing, that focuses on computing devices that "sense changes in their environment and automatically adapt and act based on these changes based on user needs and preferences"[9].

Input for user model acquisition

In order for Adaptive Systems to adapt their behaviour to the user, they must be able to acquire, and maintain, information about the user and environment. The common techniques for obtaining data to initialise user models are direct input from the user (for example, setting up a user profile) or by observing user behaviour or actions. These techniques are used in all adaptive systems, not just AH. However, according to Jameson (1999), explicit elicitation of information about the user seems to be relatively more frequent in AH than in other areas of user modelling. Use of explicit elicitation techniques was also mentioned by Brusilovsky (1996) who noted this was a measure to overcome a problem specific for AH, that observing the user's actions in hypermedia (e.g. selection of hyperlinks; time spent on particular screen) are unreliable sources of input for modelling users. Many AH systems use user feedback where possible as these are more accurate forms of data. There have been several approaches used to do this. For example, in educational systems, responses to quiz questions have been used to update aspects of the user model. Another approach is to ask users to directly provide input, either at the beginning of their use of the system by setting up a profile, or by asking for feedback on resulting adaptation to indicate whether they believe it is relevant or not. For example, Amazon (www.amazon.com), an AH recommender system that recommends products to consumers, asks users to rate items to indicate their interest in that type of item and provide additional preferences directly by setting up a user profile. A less reliable method of obtaining input to the user model is by observing certain actions performed by the user, such as clicking a link to obtain detail about a topic, as evidence from the user is interested in the topic, as in systems like POP (Höök et al., 1996).

[9] Description taken from http://en.wikipedia.org/wiki/Ubiquitous_computing, accessed 14th Jan 2006.

Methods for inferring the user model

This dimension describes the methods and techniques used to infer and abstract information about the user. For example, an AH movie recommendation system may recommend the movie *Breakfast at Tiffany's* to a user who likes *Roman Holiday*, based on the inference the user likes movies that star *Audrey Hepburn*. Recommender systems often use a collaborative filtering approach, which models similarities between users to recommend things to users based on what their peers selected or liked. Many adaptive systems use Artificial Intelligence (AI) inference techniques such as Bayesian Networks, Machine Learning, Case Based reasoning and Rules based methods. These techniques are particularly useful for dealing with input data that was noisy or involved uncertainty. Other techniques are use of user stereotypes and specifically developed rules or logic. For example, a system may use certain user actions as triggers to invoke a stereotype, inferences based on navigation actions or time spent accessing hypertext nodes.

SASY in context of Adaptive Hypertext

Table 2.1 describes SASY in terms of the dimensions of Adaptive Hypertext we have discussed above. The first column names the dimension (as have been discussed in this section), the second column lists the most common methods and approaches that have been reported in the literature. The third column elaborates on the SASY approach.

Table 2.1 – Dimensions of Adaptive Hypertext related to SASY.

Dimension of AH System	Adaptive Hypertext Literature	SASY
User model aspects used for adaptation	User data: • Goals • Tasks • Domain knowledge • Background • Hyperspace experience • Preferences • Interests • Personality type • Cognitive factors • Learning styles Environment factors: • Location • Hardware/software platform	SASY is a generic AH system. As such the user model is not geared toward a specific user attribute type and may arbitrarily contain any user characteristics of the user or environment factors listed in the previous column. The designer of each SASY application decides what attributes are modelled, and how many attributes to model, bearing in mind large user models may become difficult for the end user to comprehend and scrutinise. In SASY, the user model is a store of variables in a flat structure. For each variable, the user model also contains evidence of how attributes were inferred and metadata describing the variable.

Dimension of AH System	Adaptive Hypertext Literature	SASY
	• Usage data	
Input for user model acquisition	Direct user input User observation and inference	SASY uses a combination of directly elicited and inferred characteristics to build an initial profile. For example, user actions (e.g. clicking an HTML element, page access, answering a quiz question) may also be used to update the user model.
Methods for inferring the user model	Artificial Intelligence (AI) inference techniques: • Bayesian Networks • Machine Learning • Case Based reasoning • Rules based methods User stereotypes Specifically developed rules Collaborative filtering.	SASY applies simple user stereotypes based on user input to build an initial profile. Specific rules, as defined by the author, are used to update the user model based on user actions as described above. SASY rules could be used to infer and apply the most appropriate stereotype, so there is some support for stereotypes.

2.2 Adaptive Hypertext Methods and Techniques for exploiting the user model

In this section, we describe the common methods and techniques that have been used in AH systems to adapt the hypermedia to users, based on Brusilovsky's taxonomy (1996, 2001). Brusilovsky divides adaptation in to two categories:

- *Adaptive presentation* - the adaptation of *content* of a hypermedia node (e.g. document). The term *content* includes text, images, hyperlinks or any hypermedia element(s).
- *Adaptive navigation* - the adaptation of *hyperlinks* between or within hypermedia nodes to provide navigational support.

Table 2.2 and Table 2.3 describe the adaptive presentation and navigation techniques as described by Brusilovsky. The third column describes to what extent the technique is supported by SASY. As argued in Chapter 1, the mechanism used to generate personalised hypertext is not the focus of the innovation in this thesis. Rather standard adaptive hypertext techniques are employed as a basis for exploring scrutability of the personalisation. Overall, SASY supports basic techniques of AH as commonly applied not only in the literature but also in commercial systems, as discussed in the next section.

Table 2.2 -Adaptive Presentation Techniques

Technique	Description	SASY
Inserting/removing fragments	Hypertext pages contain adaptive fragments, that are conditionally displayed to users based on their user model. For example, a tutorial about programming loops in Java in might include an additional fragment for students that know C++, describing the difference between loops in Java and C++. The fragment would only be shown to students who know C++. For example, (De Bra et al., 2002; Kay & Kummerfeld, 1994).	This is the main technique used by SASY.
Sorting and ordering fragments	As above, fragments are displayed conditionally to users based on their user model. The order in which fragments are displayed is also based on the user model. For example, most relevant fragments are displayed first. Another possibility is that a page is made up of cascading page fragments, with the top level fragment most visible and others in the background.	Not supported by SASY.
Dimming fragments	This technique is similar to inserting/removing fragments but rather than removing a less relevant page fragment, it is displayed in a dimmer colour or less noticeable font on the page so that it is less obvious to the user. The idea is that the most relevant information on the page is most visible.	Not supported by SASY. Whilst it might be implemented through inserting/removing fragments, this is not elegant.
Fisheye Views	Similar to dimming fragments, but rather than dimming less relevant content on a page, it is reduced in font size. For example, (Tsandilas & Schraefel, 2004).	Not supported by SASY.
Stretchtext	Fragments on the page may be expanded and collapsed by the user. For example, by clicking the mouse over a word, the page content expands to include the word's definition. The page is presented to the user in a state where the most appropriate fragments are expanded and the less relevant fragments collapsed. The user may interact with the system to expand or collapse fragments, to view additional information or hide details perceived as irrelevant. In addition, the users interaction could be used as evidence to update the user model. For example, (Höök et al.1996).	Not supported by SASY. However, SASY does allow user actions, such as a mouse click, to update the user model

Technique	Description	SASY
Adaptive Multimedia and Modality	Adaptive Hypermedia systems contain information in various media including text, video, audio (music and speech). Users may have a preference as to which media they prefer, or their usage patterns may indicate that the user may benefit from or prefer a certain media type.	SASY could model the preferred of optimal media for each user and adapt its content presentation based on this. In fact, the user could be allowed to explore content freely and infer the preferred media type based on the user's actions.
Natural language adaptation	This classification is set aside for systems that use natural language technology as a foundation rather than the above fragment based approach.	Not supported by SASY. Whilst the other techniques not currently supported by SASY might be added to SASY with some additional development, Natural Language adaptation poses a completely different set of challenges that would require a quite different approach for adaptation.

Table 2.3 -Adaptive Navigation Techniques

Technique	Description	SASY
Direct Guidance	The system determines the "best" next hyperspace node for the user to visit, based on the user model and the context of the current hyperspace node the user is viewing. The system directly guides the user through hyperspace.	Supported by SASY.
Sorting	The hyperlinks presented to the user are sorted in order of relevance to the user based on their user model. For example, the user is presented a list of hyperlinks with the most relevant link at the top of the list.	Not supported by SASY.
Hiding/Dimming	The system hides or dims hyperlinks which it deems as inappropriate for the user to follow based on their user model. To mark a link as not appropriate, the system may either completely remove the navigational element from the page, disable its hyperlink property or make the link less visible to the user by dimming its appearance.	Not supported by SASY.
Annotation	The user is provided additional information about a hyperlink's state, relevance and importance to the user. The annotation helps the user decide whether to traverse the hyperlink. The additional information may be presented in many forms, for example text comments, use of images or icons, use of colours, font styles and sizes.	Supported by SASY. SASY follows the ELM-ART system by using a traffic light metaphor for colour coding hyperlinks to describe their relevance to the user. The annotated links may be used on all pages, including the SASY course map (table of contents).
Map Annotation	These techniques include all techniques which adapt the form and structure of hyperspace maps. The above techniques can also be applied to annotation in map structures.	Supported by SASY to a degree. SASY includes a Course Map that displays an annotated map of contents.

2.3 Adaptive Hypertext Case-Studies

The field of Adaptive Hypertext encompasses a large a diverse body of research, with research continuing into topics such as: user profiling and modelling (Kobsa, 2001; Diaz et al., 2004), group modelling applications (Hansen & McCalla, 2003; Farzan et al., 2005; Freyne & Smyth, 2004), application areas such as personalised TV (Baudisch et al., 2002), learning (Dolog et al., 2004), practical issues such as privacy (Kobsa & Cranor (Eds.), 2005), evaluation (Weibelzahl et al. (Eds.), 2005), frameworks and architectures (Koch, 2002; Vrieze et al., 2005), and many others some of which are discussed in (De Bra et al., 2004b). In this section, we have selected just some of the influential systems that are closely related to SASY. This illustrates the type of adaptivity we aim to make scrutable with SASY. The case studies here describe the adaptive hypertext functionality. Scrutability is discussed later in this chapter. For a more comprehensive review of AH systems, refer to Brusilovsky (1996, 2001).

2.3.1 PT

PT (Kay & Kummerfeld, 1994) tailors a hypertext that teaches the C programming language to the individual user. The initial systems that we explored through this thesis (the Tutor series of systems), were a direct extension of PT.

To use the adaptive course, users first complete an online questionnaire to elicit details about their background and knowledge of programming. Based on their knowledge, the presentation of the C course is tailored according to their user model by PT. During interaction with the course, the user completes tasks to test their knowledge of the material. This in turn updates their user model and steers the personalisation. For example, poor performance might require the user to repeat sections of the course, possibly with different examples.

The customisation is performed using a pre-processor for the C programming language. A user page request invokes a CGI (Common Gateway Interface) script that first creates a header document by translating user model attributes into variables declarations using #define. The course pages contain additional text fragments that are conditionally displayed to the user based on logic expressions involving user model attributes, for example, their knowledge of Pascal. Logic expressions in pages are C pre-processor directives. To render a page to the user, the generated header and page text are fed through the C pre-processor. The resulting page is a customised HTML document. For example, consider a raw course page such as the following:

```
A C program is a collection of functions. One must be called main and this runs first.
#if PASCAL > 0  more than minimal Pascal knowledge?
As a Pascal programmer you should note that the line main() does not have a semicolon at the end of it.
#endif
```

To render the page, PT first consults the user model and creates the header lines based on the values of user model attributes. For example, the user might have a level 3 knowledge of Pascal and no knowledge about BASIC. This would be represented as the header:

```
#define PASCAL 3
#define BASIC 0
```

Then both the header and page are passed through the C pre-processor. The resulting text is:

```
A C program is a collection of functions. One must be called main and this runs first.
As a Pascal programmer you should note that the line main() does not have a semicolon at the end of it.
```

Notice that the second line has been included in the customised text because the user model declared the user had the required knowledge level of Pascal. This type of adaptivity is called conditional inserting/removing of fragments, as defined earlier in this chapter.

Tutor 1 (Czarkowski, 1998) was in part aimed at extending this approach to provide scrutability and explore XML (eXtensible Mark-up Language) as the underling authoring language rather than text files with pre-processor commands, as used by PT. SASY also uses XML to model the user model and content definitions. Unlike PT, SASY is a generic framework for adaptive hypertext, providing a standard graphical user interface to render adaptive hypertext pages. In addition, SASY provides adaptive navigational features (direct guidance, link annotation) that are not provided by PT.

2.3.2 AHA!

Unlike many systems reported in the AH literature, AHA! is a generic framework for the creation of adaptive hypertext documents, rather than a system geared toward a specific domain or purpose. It is similar to the type of adaptive hypertext framework provided by SASY, hence it is discussed here.

The first version of AHA! (De Bra & Calvi, 1997) used a similar approach to PT and supported the same type of adaptive functionality. It also used the C pre-processor to customise HTML documents to a user, based on their user model. AHA! version 2.0 (De Bra et al., 2002, 2003a; 2003b) changed the system architecture to one utilising Java servlets, and used XHTML for adding meta-data to content to describe its adaptive behaviour. Authoring tools were developed to assist in the authoring of the adaptive hypertext material.

It works in a manner similar to PT and SASY. The author first creates a set of concepts, which define the attributes that are stored in the user model. The author also creates content pages that contain regular and adaptive content. Both require the author to create XML documents, either manually or through the use of an special purpose editor.

When a user logs in to AHA!, a user model is instantiated for them based on the user model definition. When a user accesses a page, rules, that have been defined by the author, are executed to update the user model. For example, in a teaching application, rules could be used to declare an increase or decrease of the probability that the user has knowledge of a concept(s) when they visit a page. In addition, there are some user model attributes that have special meaning for the AHA! system, like *visited*, which is used to determine how to render hyperlinks to provide adaptive navigation support, as is discussed below.

To control the presentation of adaptive content, the author defines rules in content pages that control (1) the inclusion of content fragments on a page, (2) the conditional display of whole pages. These rules are logic expressions involving user model attributes. When a user requests to view a page, the embedded rules are evaluated against that user's user model. If a rule for a content fragment evaluates to true, the fragment is included in the user's view of the page. Otherwise it is left out.

To illustrate, we repeat the same example as above, using the AHA! syntax in a teaching application:

A C program is a collection of functions. One must be called main and this runs first.
<if expr="Pascal.knowledge > 0" >
<block>
As a Pascal programmer you should note that the line main() does not have a semicolon at the end of it.
</block>
</if>

The conditional fragment "As a Pascal programmer you should note that the line main() does not have a semicolon at the end of it" is shown to the user if their user model attribute *Pascal.knowledge* has a value greater than zero. The main difference from PT is the features of the framework that allow the author to define how user model attributes are updated when pages are accessed. For example, using XHTML tags, the author can define that if the user accesses a page about beer, they might also be interested in reading about chocolate.

AHA! also provides adaptive navigation support. There are two forms discussed in De Bra et al. (2002). Firstly, navigational elements can be marked up as conditional fragments, meaning they are included/removed conditionally based on the user model. Secondly, AHA! uses hyperlink annotation. If the condition defined for a page evaluates to true, all links to that page are displayed as hyperlinks (e.g. blue font). Otherwise, links are shown in black font and are not underlined, so that they are not immediately noticeable as links. This discourages access to pages that present concepts the user is not ready to read, but it does not completely stop access to the page.

The SASY and AHA! features are similar, differing in implementation detail. Essentially, both provide similar adaptive presentation and navigation techniques. AHA! and SASY were developed

independently and in parallel, with the scrutability goal driving work on SASY differentiating its direction from that of AHA! Recent work with AHA! has been towards extending the AH functionality in AHA! version 3, creation of authoring tools (De Bra et al., 2002b; 2003), integration with other systems such as INTERBOOK (De Bra et al., 2003c), CLAROLINE (Arteaga et at., 2004).

2.3.3 ELM-ART

ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001), an adaptive hypertext system that teaches the LISP programming language, uses a completely different style of adaptivity from PT and AHA! Since initial development in 1996, it has gone through several versions. Here we describe the version presented in (Weber, Brusilovsky, 2001). ELM-ART was a large influence for SASY, particularly, for its user interface and functionality: Notes editor, Annotated Course Map and hyperlinks.

It uses a form of an overlay user model and episodic student modelling to provide adaptive navigational support, course sequencing and individual diagnosis of student solutions to quiz questions. ELM-ART represents knowledge about domain concepts to be learned in terms of a hierarchical, conceptual network. This is overlaid with a multi-level user model that captures a visited state (whether a page has been visited), learned state (which exercises relating to the concept have been worked on or solved), inferred state (whether a concept might already be known based on related known concepts), known state (whether a concept is known, having been set so explicitly by the user or their collaborators). The user is able to access and modify their user model, but this is discussed later in this chapter.

The user model is used to provide adaptive annotation of hyperlinks in ELM-ART. Links are visually annotated according to a traffic light metaphor. A green ball (next to a link) means the system suggests the user follows this link next, as the concepts on that page are ready to be learnt by the user. A red ball means the concepts are not ready to be learnt because the student does not know one or more pre-requisite concepts yet. A white ball indicates the concept is already known by the student. A yellow ball means the user has not fully learned the concept. This metaphor for hyperlink annotation has been applied to SASY.

In addition, the system uses direct guidance, with a *Next topic* button, to help the user navigate through the course. This is a link to the next page that should be read by the user, computed by the system based on the next concept that needs to be learnt to satisfy later pre-requisites.

A typical lesson page in ELM-ART is shown in Figure 2.1. The annotated course map is shown on the left of the image. Here the user has just started the Datatypes topic. The first concept to be learnt is Atoms. Notice the green ball next to the Atoms link in the coursemap, indicating the page is ready to be accessed by the user. The other links are red, as the learner is not ready to read these pages. Notice also

the arrows at the bottom of the page. These are hyperlinks, directly guiding the learner to the next most appropriate page.

ELM-ART is the basis for a commercial adaptive course authoring system, NetCoach (http://www.net-coach.de). The adaptive navigation techniques of direct guidance and annotation of colour coding through hyperlinks have been applied to SASY.

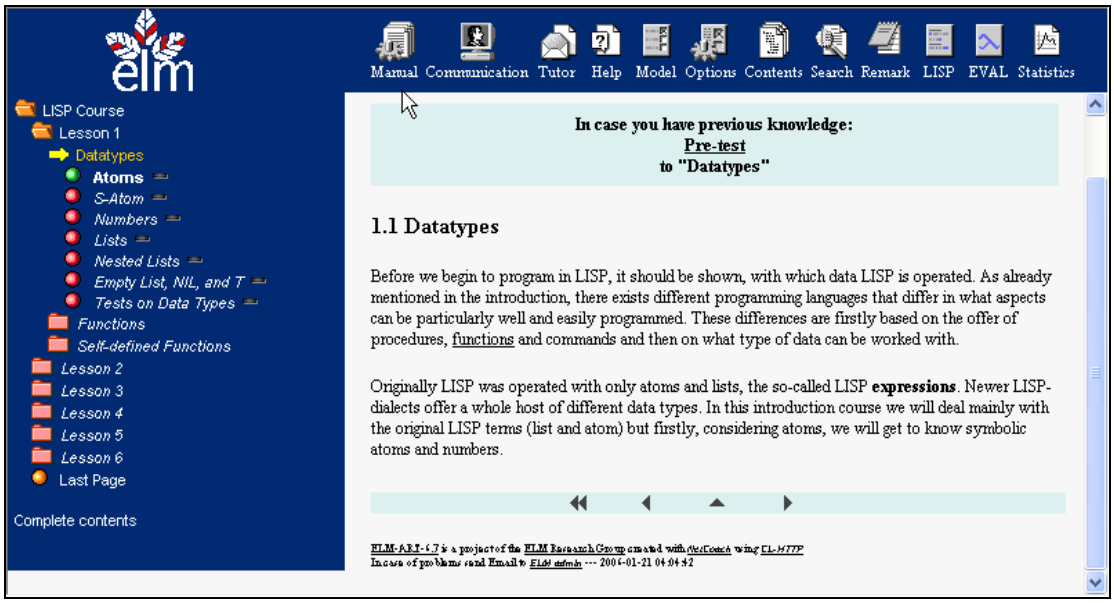


Figure 2.1 - A typical lesson page in ELM-ART.

2.3.4 Commercial Adaptive Hypertext Systems

This section provide an brief cross-section of the type of products available for commercial adaptive hypertext. Several commercial vendors offer products for adding personalisation to hypertext, based on conditional inclusion of fragments: ATG Personalisation Server, BEA WebLogic, and IBM Websphere Personalisation. This is the type of personalisation we aim to make scrutable with SASY. We describe these three vendors as their approach to content adaptation is quite similar to SASY and information about their architecture and approach is readily available from their website. There are other well known commercial vendors that use a collaborative filtering approach to content adaptation such as BroadVision (http://www.broadvision.com) and NetPerceptions (http://www.netperceptions.com), a product based on the approach developed by the GroupLens research group (http://www.cs.umn.edu/Research/GroupLens).

ATG Personalisation Server

ATG Dynamo Personalisation Server (http://www.atg.com) provides features to allow web site creators to target content to users, or user groups. ATG was one of the first vendors to provide this type of commercial product.

The author first defines what attributes about their users are to be modelled in their profile. These can be simple values, for example, gender, age, or a hierarchical tree structure of attributes. Next the author defines rules, based on attributes in the user model, and assigns rules to content. Rules can be created through a graphical user interface or through code, using custom SGML code (Standard Generalized Mark-up Language). Authors also have to develop code for populating the user model. However, as the user model data is stored in a database, it could also be the case that the user model is provided by another source, such as a marketing database. For example, if a customer has bought three expensive cars at a car dealership within a short time, the dealership might flag the customer as a big spender in their customer database.

When a user browses a web page, the rules embedded in the page are processed against their user model to determine what content is displayed and what is omitted. For example, Figure 2.2 defines a rule set to match users who have an attribute *BigSpenders* set to true in their user model, and *interests* attributes includes *cars*.

```
<ruleset>
  <accepts>
    <rule op="and">
      <rule op="eq">
        <valueof bean="Profile.BigSpenders">
        <valueof constant="true">
      </rule>
      <rule op="includes">
        <valueof bean="Profile.interests">
        <valueof constant="cars">
      </rule>
  </rule>
</accepts>
</ruleset>
```

Figure 2.2 - Example rule set for an ATG page, based on documentation from the product website (http://www.atg.com).

WebLogic Personalization Server

The BEA WebLogic Personalsiation Server (http://www.bea.com) offering is similar to ATG. The BEA E-Business Control Centre is a graphical user interface tool that is aimed at business analysts and marketing professionals. It is made up of three components: BEA Campaign Manager, WebLogic Commerce Server and WebLogic Personalization Server. The Campaign Manager allows these authors to analyze customer interactions, create placeholders for content display, define customer segments to target customers with personalized content, define campaign discounts, execute and manage promotional campaigns, and monitor and report campaign results. For example, Figure 2.3 shows how one can define a customer segment to represent frequent visitors. That is, users who have visited the web site four or more times.

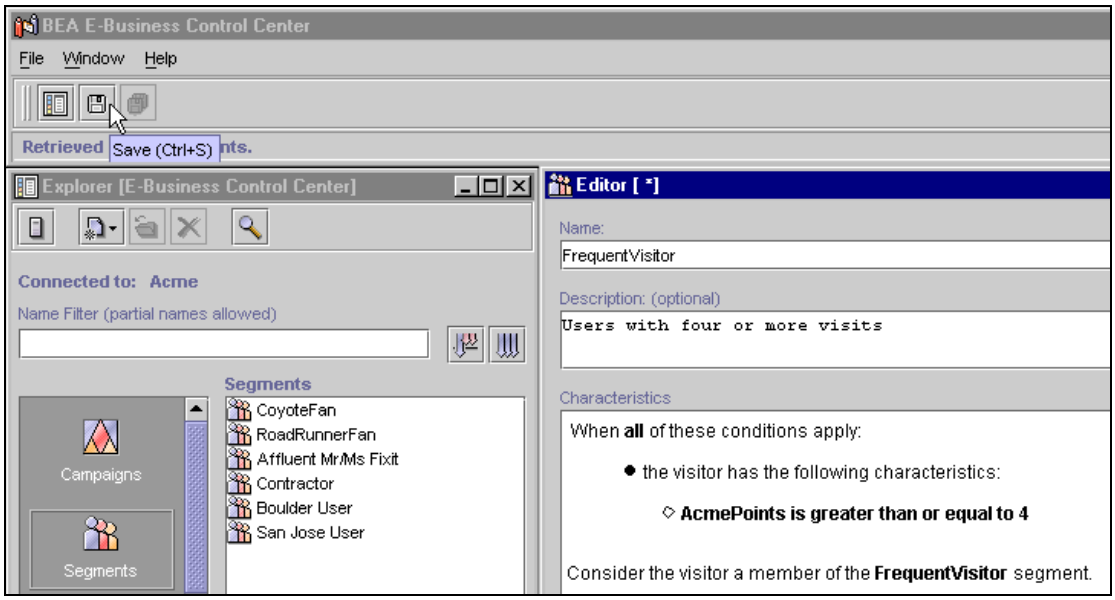


Figure 2.3 – BEA Customer Segment editor.

Image from BEA product documentation available online at http://www.bea.com.

The application server keeps track of the customers (end user's) behaviour by recording certain events such as: clicking on a link for a campaign, browsing or reading about a product, adding a product to shopping cart. The Commerce Server allows technical authors to create and deploy J2EE, web based application, based on existing templates and run out-of-the-box. Content pages are created using JSP (Java Server Pages technology). The WebLogic Personalization Server, allows users to personalise content to a user using attributes in their profile. This is done by defining rules based on profile attributes, and assigning rules to content, in a similar way to that described for the ATG product above. Rules are defined by adding JSP tags to content pages. There are also JSP tags for getting and setting user model attributes.

IBM Websphere Personalisation

IBM's product also provides web site personalisation based on rules for conditional fragment inclusion based on user profiles. The graphical interface includes: the Campaign Manager (similar to BEA's Campaign manager described above), a Rules Composer for defining personalisation conditions, shown in Figure 2.4, and a Preview Launcher, to preview the resulting personalised page. The Preview Launcher renders the content as an HTML page, as shown in Figure 2.5. Blue dots are included in the page to show slots where adaptive content would be included by the personalisation server, should the condition for the content match the user's profile. Two dots are shown in Figure 2.5: one in the first paragraph following the word *Welcome*, the second under the title *Your Company News*.

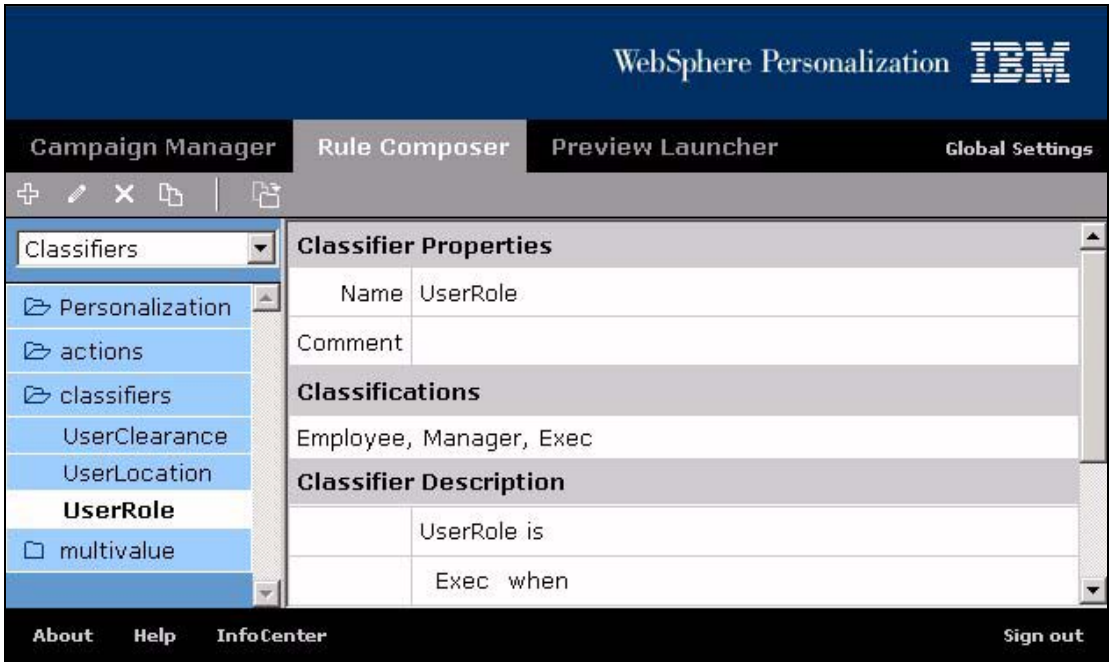


Figure 2.4 – IBM WebSphere Personalisation rules editor.

Image (and that of Figure 2.5) from product documentation available online at http://www-306.ibm.com/software/genservers/personalization/Pers4InfoCenter/en/main/guide/howto.html#intro, accessed 24 February 2006.

Figure 2.5 - Preview of personalised page in IBM WebSphere.

In addition, IBM provides recommendation-based personalisation based on collaborative filtering. This allows the delivery of personalised content based on what was shown to other users who made similar choices in the past. However, details of the functionality are not published by the vendor.

Commercial Adaptive Hypertext Systems and SASY

The commercial products described in section all aim to adapt to a user based on knowledge of their interests, characteristics, purchase history, with an overall goal of increasing online sales or improving the end-user's experience. The approach for content personalisation is the based on the same principle across these three commercial systems and SASY: conditionally including content based on the evaluation of user defined rules involving user model attributes. One reason for the popularity of this approach to personalisation is its simplicity. Interestingly, none of these three commercial products provide out-of-the-box support for the adaptive navigation techniques described in the previous section.

2.4 Openness in Adaptive Systems

This section presents work that relates to the openness of adaptive systems. The term *open* may have different meanings, as discussed by Self (1999). The *openness* aspects we focus on here relate to the user model, and how the user model is used to provide an adaptive response to the user. In particular, most relevant to the topic of this thesis, we discuss research that considers the openness and controllability or the adaptivity and associated processes.

Bull and Kay (2006) use a framework, called SMILI (Student Models that Invite the Learner In) to describe the openness aspects adaptive learning environments. However, it also applies to adaptive systems that are used for purposes other than learning. First we use this framework to show the context of the work of this thesis. Table 2.4 shows the SMILI framework, which has been re-arranged so that the aspect most relevant to this thesis is discussed last. The first set of columns show the reasons for providing openness of the user model or adaptation as:

- Increasing the correctness of data in the user model, e.g. Mr Collins (Bull & Pain, 1995).
- Facilitating collaboration by allowing users to understand each other better by accessing each others model, e.g. (Hansen & McCalla, 2003).
- Guiding navigation, e.g. Knowledge Sea II (Farzan et al.,2005), QuizGuide (Brusilovsky et al. 2004).
- Issues related to the user's right of access to data about themselves and being able to control the data. Also involves increasing trust in a system by making the adaptivity transparent (Höök et al.1996), and supporting a sense of curiosity about how the system works, or how the adaptivity might be changed to better suit the needs of the user.
- Purposes specifically related to learning:
    - Promoting learner reflection on knowledge and understanding, e.g. SIV (Li & Kay, 2005);
    - Helping learners, or their teachers to plan and/or monitor their learning based on the information available in the learner model, e.g. ViSMOD (Zapata-Rivera & Greer, 2004).;
    - Using the learner model, or part of it, as an assessment of the learner, e.g. SQL-Tutor (Mitrovic & Martin, 2002), PSAT/NMSQT (Zapata-Rivera et al., 2005).

The first two rows of Table 2.4 describe general considerations: whether the openness is central to the system's purpose, for describing open learner models, and evaluation or evidence to support the benefit of the openness aspect.

Table 2.4 -The SMILI Open Learner Modelling Framework (Bull & Kay, 2006).

Columns show the purposes of the openness of the learner model. Rows have the elements that provide the means to achieve those purposes. Critical elements for a purpose are marked with X, arguably important ones with = and others are blank. The right most column describes aspects that relate to SASY: O means the aspect is present, N means it is not. Thanks to Susan Bull and Judy Kay for allowing the table to be presented here.

Purpose / **Elements**	*Properties Description*	*Accuracy*	*Collab.*	*Nagvigate*	*Right of Control*	*Reflection*	*Plan/ Monitor*	*Assess.*	***SASY***	***SASY - Details***
i. Centrality of OLM										
ii. Evidence / evaluation										
1. Extent of model accessible	Complete	=	=	=	X	=	=		O	Full access
	Partial	X	X	X		X	X		O	
	Knowledge level	X	=	X	X	X	X	X	=	User definable, but
	Knowledge	X	=	X	X	X	X	X	O	no inference
	Difficulties	X	=	X	X	X	X	X	=	capabilities
	Misconceptions	X	=	X	X	X	X	X	=	
	Learning issues	X	=	X	X	=	=		O	Can not access other
	Preferences	X	=	X	X	=	=		O	users' models
	Other	X	=	=	X	=	=		O	
	Other users	X	X			=	=		N	
2. Presentation	Textual (i.e...)								O	
	Graphical (i.e...)								N	
	Overview	X	=	=	X	X	X	X	N	SASY presents full
	Targeted/all Details	X	X	X	X	X	X	X	O	user model, no
	All Details	=			X	=	=		O	overview provided.
	Support to use	X	X	X	X	X	X		O	Purpose of each user model attribute is explained by system.
3. Match underlying representation	Similar	=	=	=	X	=	=		O	
4. Access to uncertainty	Complete	=	=	=	=	=	=	X	O	
	Partial	X	=	=	X	=	X			
5. Role of time	Previous	=		X	X	=	X		O	Shows current model
	Current	X	X	X	X	X	X	X	O	and historical
	Future	=		X	=	=	X		=	evidence.
6. Access method	Inspectable	X	X	X	X	X	X	X	O	Full control is given
	Editable	=			X				O	to users who may

Purpose / **Elements**	*Properties* / *Description*	*Accuracy*	*Collab.*	*Nagvigate*	*Right of Control*	*Reflection*	*Plan/ Monitor*	*Assess.*	***SASY***	***SASY - Details***
	Student persuade	X			=	=		X	N	change user model
	System encourage	=				=			N	without contest.
	Negotiated	X			=	=		X	N	
7. Access initiative	System	X	=	=		=	=		N	User initiates all
comes from	User	X	=	X	X	X	X	X	O	scrutinisation
	Peer	X	X	=		=	=	=	N	activities.
	Instructor	X	=	=		=	=	X	N	
	Other	=	=	=		=	=		N	
8. Access to sources	Complete	X		=	X	=	=	X	O	Provides evidence of
of input	Partial	X	=	=	X	=	=		N	user model values.
	System	X		X	X	X	X	X	O	Evidence describes
	Self	X	=	=	X	=	=	X	O	whether model value
	Peer	X	X	=	X	=	=	X	N	was set by user
	Instructor	X		=	X	=	=	X	N	directly or inferred
	Other	X		=	X	=	=	X	N	by system.
9. Control over	Complete		X		X				=	Other users have no
accessibility (to	Partial		X		X				N	control.
others)	System				X				N	Instructor can view
	Peer		X		X		=	X	N	user model through
	Instructor				X				=	command line utility.
	Other				X			X	N	
10. Flexibility of	Complete	=	=	=	X	=	=	=	O	User can configure
access	Partial	X	X	X	X	X	X	X	O	visibility of elements
										that describe user
										model values on
										adapted pages.
11. Access to model	Complete	=		X	X	=	X		O	This dimension is the
effect on personal'n	Partial	=		X	X	=	X		O	main focus of SASY.

The next set of rows (numbered 1 through 10), identify the dimensions of openness, that allow the system's purpose for openness to be achieved.

Extent of Model Accessible - Row 1 describes how much of the user model is open to the user and which attributes of model are accessible to the user. The first aspect of this dimension is whether users have full access or partial access to user model data. By default, SASY allows users to access all aspects of their user profile. However, the SASY framework also allows the use of private user model attributes that are used by the system or author but not exposed to the end-user. So we support both full and partial access in SASY.

The second aspect of this dimension is what types of knowledge details are opened to the user, mostly applicable to adaptive educational systems. SASY is a generic adaptive hypertext system, so unlike an adaptive system designed for educational, such as ELM-ART, it does not have the ability to deeply model knowledge and misconceptions, and diagnose student errors. Several systems have used a *skill meter* to make the knowledge modelled about a user accessible to them (Papanikolaou et al, 2003; Weber, Brusilovsky, 2001; Mitrovic & Martin, 2002). For example, ELM-ART (see Figure 2.9) used a partly filled in horizontal bar to indicate the percentage of learnt material. There is also work that used a skill meter to show misconceptions (Bull & McEvoy, 2003).

The final aspect of this dimension describes openness of other user model attributes. The types of data that can be modelled about a user was discussed in the previous section, in the Jameson framework for describing adaptive hypertext. There are examples of making these attributes open to the user. For example, opening learning strategy to the user (Bull et al., 1995); level of trust (Johnson et al., 2005); usage data in the form of products that had been purchased and rated by the user that influences the user model (Amazon online retailer, http://www.amazon.com). Finally, Linton & Schaefer (2000) have explored allowing users to compare their knowledge to the combined knowledge of other groups of users (Linton & Schaefer, 2000). SASY does not distinguish between different user attributes and can model arbitrary user model attributes. SASY does not provide access to user models of other users.

Presentation – Row 2 describes how the information is opened to the user. As argued by this thesis, this dimension must cover not only the user model but also the personalisation process that describes how the user model is used. The presentation might be textual or use a graphical interface. This dimension also covers whether the user can access an overview then drill down to focus on certain aspects, and whether there is special support to help the user understand the information presented. In terms of opening the user model, SASY presents complete details textually, and also describes how attributes are used in the personalisation process. On personalised pages, SASY presents an overview of the number of items that were personalised, shows relevant aspects of the user model that influences personalisation and allows the user to drill down to see full explanations of each adaptation. This is done textually but is also annotated with colour, by highlighting personalised items on the page. Like ELM-ART, SASY also uses colour annotations for hyperlinks as described above. Another example of colour is in Knowledge Sea II (Farzan et al.,2005) which uses shades of colour to support social navigation by representing grades of what has been explored and unexplored by a group, reflecting the user model of the group. Graphical formats have been used to describe more complex user models, such as: a representation of a Bayesian network (Zapata-Rivera & Greer, 2001); a hierarchical tree structure (Kay, 1999); a conceptual graph (Dimitrova, 2003); visualisation of large user models (Kay & Uther, 2003). A limitation of the way SASY provides access to the user model it that for large user models, the profile may be difficult to comprehend because all of it is made available to the user on a single page. At the same time, useful adaptations can still be performed using small user models.

Match underlying representation – In SASY, the representation of the user model and personalisation process matches what actually occurs within the system. However, this may not always be appropriate. For example, special graphical formats have been used to present user model details to young users (i.e. children) in ways they can understand (Bull et al., 2005b; Bull & McKay, 2004).

Access to uncertainty – Row 4 describes whether the user can access information about the certainty of information in the user model, for example, the probability that a learner knows a concept (Corbett & Anderson, 1995; Corbett & Bhatnagar, 1997; Kay, 1999). SASY provides a list events that influenced the value of user model attributes as evidence. If probabilities are used to define attributes in the user model, the user has direct access to this information.

Role of time – This dimension describes whether users can look back so see previous user model states and predict future states. This is the case for knowledge states in the work of (Bull et al., 1995). SASY, like some other systems (Kay, 1999; Kay & Uther, 2003; Kay & Lum, 2005), allows the user to view evidence of attributes in the model. This describes what change was applied to the user model and when it occurred. Although, the user can not examine the state of the user model at a particular point in the past, they are able to change their user model to represent a previous state.

Access method – Row 6 describes the type of access users have over their user model. Users may be given access to elicit initial user preferences and later be allowed to inspect and edit their user model (Bull & McEvoy, 2003; Kay, 1999, Weber & Brusilovsky, 2001), as in SASY. Other systems require the user to negotiate the user model values to justify the change (Bull & Pain, 1995; Dimitrova, 2003); or the user can influence the model by undergoing assessment of their knowledge (Weber & Brusilovsky, 2001; Brna et al., 1999).

Access initiative – This describes who initiates access to the user model. Scrutability might be sought by the system, the user (as in the examples above), by peers to allow students to compare their progress to others (Dufresne & Hudon, 2002; Tongchai & Brna, 2005), to externalise user models to instructors in distance learning situations (Mazza & Dimitrova, 2003). In SASY it is always the user who initiates access.

Access to sources of input – This dimension covers what information is provided to the user to describe where information in their user model came from. For example, explaining whether it was inferred by the system or provided by the user directly, as in SASY, or perhaps from an instructor, peer or some other source.

Control over accessibility (to others) – This dimension describes the level of control the user has over allowing their user model to be exposed to others. For example, in UMPTEEN users can open their learner model to their peers, instructors in named or anonymous form, and view individual peer models

to which they have been given access, and the group model (Bull et al., 2005). This dimension does not apply to SASY as users can only scrutinise their own user models and adapted system response.

Flexibility of access – This dimension describes whether the presentation of the information is flexible, whether it can be presented in different formats, differing levels of detail or accessed through different methods. For example, Subtraction Master, discussed in (Bull and Kay, 2006), is an educational system that presents a different view of the learner model to the users (children) that what is shown to their teachers. The student's view is simplified and presented in a form children are able to comprehend. In SASY, the user can configure how scrutinisation information is presented. The user can choose whether scrutinisation is shown on every page by default, or not at all.

Access to model effect on personalisation - The last row is the dimension that is the topic of this thesis. It describes whether users are able to determine the effect of the user model on how the interaction is personalised for them. Whether the personalisation is explained to them and what level of detail is provided in the explanation. Here we describe non-AH examples, we devote the next section to AH examples, since they are most relevant to this thesis.

Firstly, it is important to distinguish between implicit and explicit ways in which the user can determine the effect of the user model on personalisation.

Implicit forms are those that do not actually provide an explanation of the personalisation to the user, but the user is still able to figure out some of the details. For example, consider an adaptive recommender system that personalises a television viewing schedule based on users interests. A user might guess that if they specify an interest in the genre of music, that the system might include music programs in their personalised view. A user could also remove/add interest in a genre and observe the effect on personalisation.

On another level, the system might provide explanations of how it inferred beliefs about the user, which might also provide some information about how personalisation is affected. For example, Mr Collins (Bull & Pain, 1995) is a system for foreign language learning. It exposes its learner model to the user and allows them to query the system to see how it inferred its beliefs about the user. The user can then modify the learner model through a process of negotiation with the system. The information provided by the system describes how it inferred the user's knowledge, which suggests how personalisation is affected by the current state of the learner model. This is also the case for another system, StyLE-OLM; (Dimitrova, 2003).

In this thesis, we focus on explicit explanations for adaptation. This means the user is told exactly why a personalisation occurred, rather than providing hints. As a result, the user is also in a better position to understand the implications of changing the user model. A recent example of this same approach is the Intelligence Office System (IOS) (Cheverst et al., 2005). IOS learns a given user's behaviour in an

office environment and builds up 'IF-THEN' rules to describe the behaviour. It then proactively adjusts physical elements in the office environment (i.e. fan, desk lamp and heater) to suit the user by executing the rules. For example, it may learn the user typically turns on their office fan if the temperature increases beyond a threshold. It then monitors the temperature through hardware sensors and automatically turns on the user's fan if it gets too hot, saving the user from having to do it themselves. A key goal of the research is to allow the user to feel as though they can understand and control the adaptive behaviour of the system. To achieve this, it enables the user to scrutinise and possibly over-ride the 'IF-THEN' rules held in their user model.

The main IOS interface is shown in Figure 2.6. The left part of the screen allows users to switch on/off the office fan, heater and desk lamp. The right hand side of the screen shows the current sensor readings, and provides buttons through which the user can view and edit the data collected about their behaviour and preferences. When IOS senses a set of environmental conditions that trigger a rule to be fired, and if the user has selected that they want to be prompted before changes occur, IOS changes the interface to appear as in Figure 2.7. Notice at the bottom of this figure the system message "I believe you want to turn the fan off (confidence level: High)" and the button to the left of that labelled "WHY??". Clicking this button changes the interface to appear as in Figure 2.8. Here IOS describes the rule "If Temp = cold then Fan → off" that fired. The user can further scrutinise by clicking the rule. This brings up evidence of their behaviour that led to the rule being inferred: essentially, instances of the user's behaviour in the past that match this rule. The user can then view other associated rules for the device (fan, heater of lamp). Users can also tell the system not to use the rule again (see top left button in Figure 2.8), add new rules and modify the boundary thresholds (for example the temperature range that defines *too hot*).

Cheverst et al. (2005) reported in their qualitative study, with 30 participants, that 90% expressed the need for control over the systems adaptive behaviour and 66% wanted access to basic and in-depth explanations. The latter meant the ability to look at the rules which had been generated by the system. In addition, 69% thought the textual descriptions of rules, as at the bottom of Figure 2.7 were clear and easy to understand.

There are a number of parallels between IOS and our work with SASY. Firstly, the types and depth of scrutinisation is similar. In IOS, personalisation causes a device (e.g. heater) to be turned on or off; in SASY a content fragment is included or removed to create the personalised view of a page, but the use of IF-THEN logic expressions to define personalisation is common to both. Another commonality is that the user can see an explanation of why the personalisation occurred based on inferred information stored in the user model, and view evidence of why the beliefs were inferred.

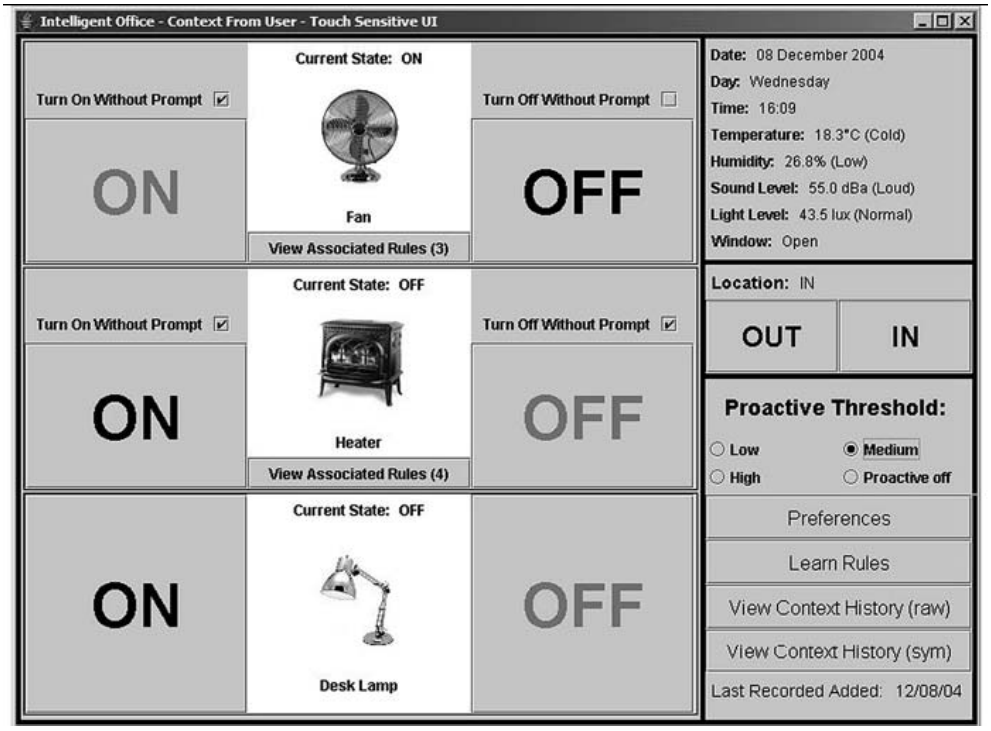


Figure 2.6 - Intelligence Office System main interface.

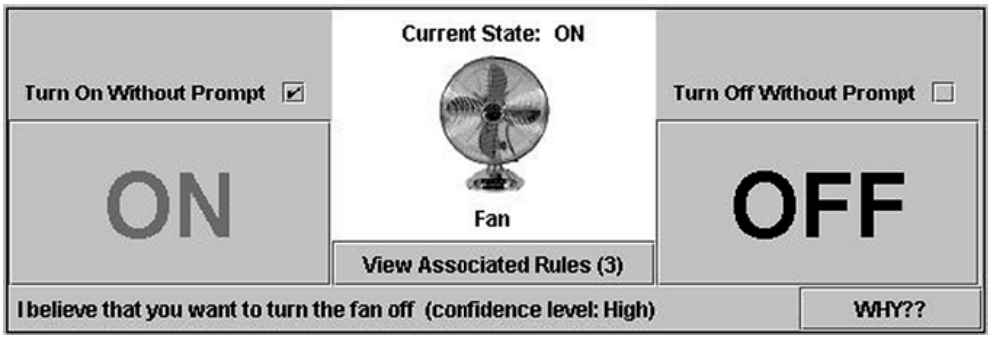


Figure 2.7 - IOS explanation of proactive action.

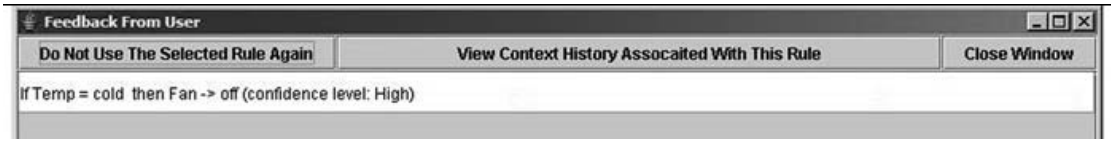


Figure 2.8 - Scrutinising the personalisation in IOS.

There are also a few key differences. The first is a subtle but important difference relating to the system's purpose. Users access the IOS interface to turn devices on or off, or to scrutinise the personalisation that proactively performs this function on their behalf. On the other hand, users access SASY to access information, for a range of reasons. It could be argued that with SASY, the desire to scrutinise the personalisation is more distant from the user's overall goal for using the system compared to IOS. This, in turn, adds difficulty for the interface design for the scrutability components

in SASY. Another subtle point is that in IOS, the factors that influence personalisation are fixed sensors, while the personalisation rules are inferred. In SASY, attributes are inferred about the user while the rules are fixed. For example, SASY might infer the user has an interest in cars based on their actions using the system. Finally, SASY adds scrutability to a generic adaptive hypertext framework for creating various, domain independent, adaptive hypertexts.

Another system that provides explicit explanations for adaptation is location aware, mobile shopping guide (Bohnenberger et al., 2005). The system was developed to run on a PDA (personal digital assistant), that a shopper would use to guide them through a shopping mall to help them achieve their objective: to purchase certain goods in a time critical situation. The shopper specifies at the beginning their interests in particular types of products. The system computes a policy to direct the shopper to the next store based on: their current location, the products purchased so far, the amount of time remaining. The initial version of the system computed a policy and guided the user from store to store, indicating the direction in which to walk to arrive at the next store and the time remaining for the shopping task. However, it did not provide any explanation of how it generated the policy, nor did it explain the implications if the user deviated from the route mapped out by the system. An initial evaluation found that users were hesitant to deviate from the recommendations, even when they had good reason to do so, because they had no visibility into the recommendation process and did not know how this would impact their overall goal. That is, they could not scrutinise the recommendations to understand why they were made or investigate alternatives.

The next version of the system was enhanced to include visual explanations. In this version, the system presented up to three alternative routes through the shopping mall, indicating the time each route would take and the stores that would be visited on the way (and the order). The system also indicated which was the optimal route. The explanation allowed the user to compare route duration against their actual time remaining to complete the shopping task and see what stores would be visited in each alternative. This provided some insight into the implications of choosing an alternative.

The new version also considered when to explicitly present explanations to the user and when to make it optional for users to access this information: when the difference between the overall utility of the recommended option over the next option was small, the user had to explicitly request to view the explanation; when the difference was large, the system automatically presented it. The rationale was that it is important for users to view the explanations to deter them from deviating from the recommended route, in situations where doing so might drastically impact their goal. The follow up evaluation found that users varied in their attitudes toward the usefulness of the explanations, the timing of explanations and user control over them. Specifically, the study reported that: (1) under situations of high time pressure, the use of explanations imposes additional load on the user that may lead to poorer performance in terms of the overall objective to purchase items in a limited time; (2) the explanations helped clarify the reasons for the system's recommendations; (3) some users would prefer more control over the explanation presentation (i.e. controlling whether it is presented automatically).

The study concluded that on one hand users might not want detailed explanations of the personalisation, but, on occasion, might wish to user their judgement to override the systems recommendation and doing so will require more information from the system about how it arrived at a particular recommendation (e.g. background, system capabilities, limitations).

2.4.1 Scrutable Personalisation in Adaptive Hypertext Systems

In this section we describe research that has explored the scrutinisation of personalisation in adaptive hypertext systems. We present the systems that most influenced our work and research that is most closely related.

ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001), the adaptive hypertext LISP course application described earlier in this chapter, was a major influence on SASY in terms of the learning environment is provides. ELM-ART allows users to inspect and update their user model. An example of the interface to the user model is shown in Figure 2.9. The left side of the page shows knowledge concepts that are modelled, the middle columns indicate the student's progress through the course, the right hand side of the page allows the user to mark concepts as known and update their user model. This form of user model viewer is sometimes called a *skill-meter*. Unlike SASY, the user can not see how the user model concepts affect the personalisation of content.

Another important piece of early work that explored transparency of adaptation was the POP prototype system, part of the PUSH (Plan and User Sensitive Help) project (Höök et al.1996). This adaptive hypertext system personalised the presentation of online technical project documentation. Pages were personalised using the stretchtext technique (described earlier in this chapter) where page sections are expanded and collapsed, and POP also conditionally displayed hyperlinks and hotwords that allowed users to drill down for more detailed information. When a user first accesses the system, they select their key task. This initiates a stereotype that controls how the first page is presented: which sections are expended or collapsed initially. The system updates the user model to infer any changes to the users task as users manipulate pages: expanding and collapsing sections, clicking on hotwords and typing free form search queries.

Users can scrutinise POP in two ways. Firstly, the system explains its inferences about the user by displaying what it believes is their current task on the hypertext page. For example, it might display the following message "This answer is assuming that your current information seeking task is Project Planning". Secondly, if the user disagrees with the system's inference they can change the task, thereby updating their user model. In SASY, we are interested in more fine grain and detailed explanations of each personalisation. However, SASY also indicates beliefs about users that influence personalisation on each page.

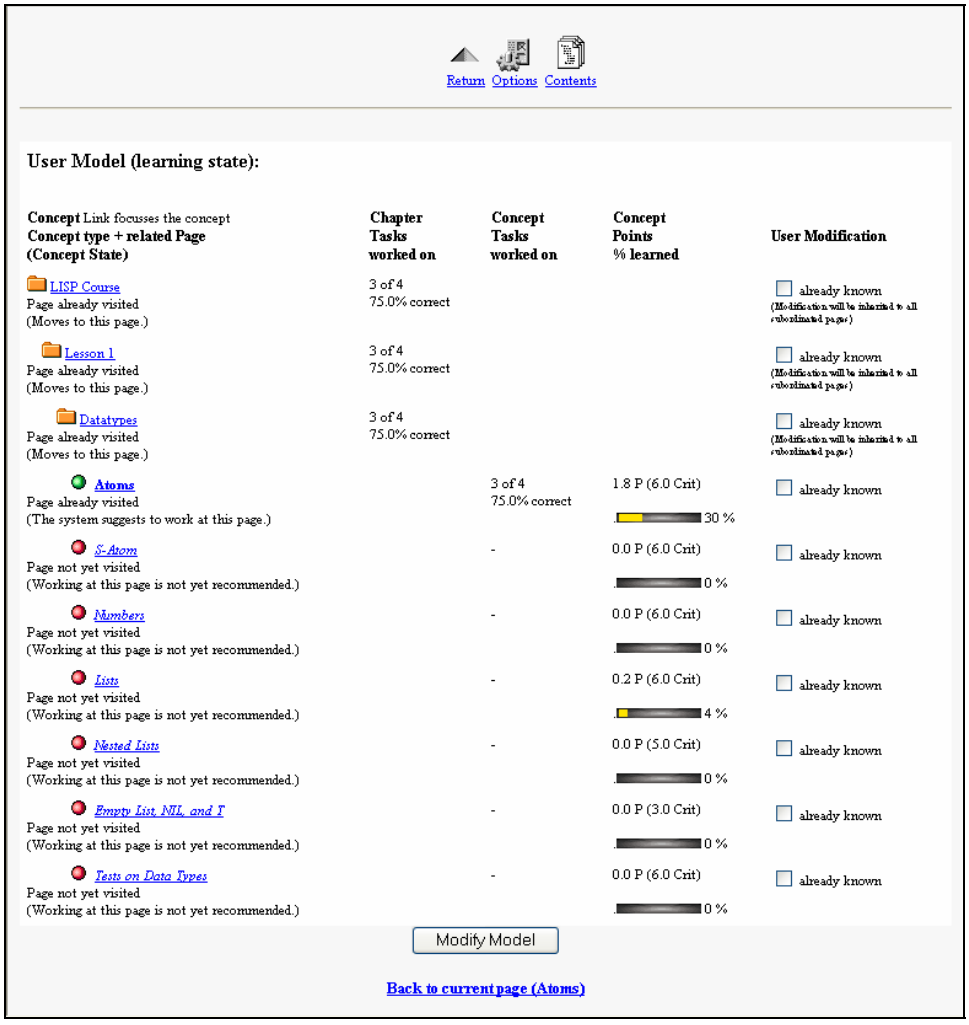


Figure 2.9 - Inspecting user model in ELM-ART.

There has also been research to explore the user's ability to control personalisation. However, the type of control is not consistent in the literature. One type of control is allowing the user to specify whether personalisation is performed at all. For example, in HYLITE+ (Bontcheva et al., 2005) users can control whether or not the system updates, uses their user model and whether the system performs adaptation at all. Similarly, in Jameson & Schwarzkopf (2002) users can control when the adaptation occurs. The evaluations of both systems have reported that users desire this type of control. Another type of control is allowing the user to influence or dictate how the personalisation works, how it generates its response: this is the type of control we focus on in this thesis and discuss next.

In the adaptive hypertext literature presented in this section so far, the user has insights into how the personalisation works, but not an actual explanation of how their user model is used for personalisation. The following two works give the user a deeper explanation of the personalisation that has taken place. Unlike SASY, they are pure recommender systems so their technique for personalising content to the user is different.

Jameson (2005) describes an adaptive hypertext system that presents details of a conference to potential attendees. The feature of the web site that we discuss here is the adaptive *hotlist*. The *hotlist* is a book-marking tool that allows the potential conference attendee to create a list of conference events they are interested in attending. The user can manage the *hotlist* by adding and removing events, as per their interests. The system also makes personalised recommendations for events the user might be interested in, based on inferences of the user's interests, by adaptively adding recommended events in the *hotlist*. The user may accept or reject recommendations. An example of the interface is shown in Figure 2.10. The first part of the screen is called the *hotlist*. Sessions that are recommended by the system are displayed in red font. For example, in the third row of Figure 2.10, the session titled *Tailoring privacy to the user's needs* is shown in red. Clicking a *View Session* link shows the full description details about the session on the bottom half of the page, that includes an explanation as to how an event fits the user's interests. The explanation lists the concepts that relate to the event and the user's inferred interest level in each, indicated by plus sign(s) (to indicate interest) or a minus sign(s) (no interest). The scale ranges from ----- (5 minus signs indicate lowest interest level) to +++++ (high interest level). For example, the second last line in Figure 2.10 shows the description "Information retrieval (++), E-commerce (++)", meaning the event relates to both *Information retrieval* and *E-commerce*. Also, the system has inferred the user has some interest in *Information retrieval*, 2 plus signs, and even more interest in *E-commerce*, indicated by 3 plus signs. In this version of the conference website, the user also has control over when the personalisation is performed: clicking the *Update recommendations* button (top right of screen in Figure 2.10) updates the *hotlist* to possibly include new event recommendations, determined by the system. The user can manually control the *hotlist*, by selecting check boxes for events in the *hotlist* to accept/reject new recommended events, or remove existing events from the list. Clicking the *Execute changes* button applies the selected changes to the *hotlist*.

There were two main studies of this system discussed by Jameson (2005). The first, explored user attitudes to control over when personalisation occurred (Jameson, 2002). The experiment used three system variants[10]: (1) the user could control when system recommendations were applied to the *hotlist* by clicking the Update button (as described above); (2) the system *automatically* updated the *hotlist* whenever the user added, removed, accepted, rejected an event, so there was no Update butto; (3) the system made no recommendations at all. The study found users differed in their preferred level of control. Some preferred control over when the *hotlist* was updated with recommendations, others preferred it was updated automatically by the system. It concluded different users might want different levels of control, in different situations.

[10] Note: this experiment was conducted with a previous version of the system, refer to (Jameson, 2005) and (Jameson, 2002) for details.

The second study, not currently published, but discussed in (Jameson, 2005), explored whether the recommendation explanations improved user comprehensibility of how the system generated personalised recommendations. The study used two system variants: (1) with explanations; (2) without explanations. The study found users who were provided with explanations had more insight into what the system took in to account to make recommendations. The study also found most users found the explanations "somewhat useful" or "useful to a small extent".

In the current website, users can not update their user profile to change their inferred interest level(s). Jameson states that in both evaluations, the ability to do so was one of the most-desired missing features.

Figure 2.10 – Screen shot of recommended conference session, showing an explanation of the adaptation.

Another recommender system that allows user to scrutinise, but in a different way, is the online retailer Amazon Inc. (http://www.amazon.com). Amazon builds a profile of each user's interests based on their history of purchased and browsed items. The user can also volunteer their interests by setting up a profile and rate products to provide additional information in to the user model. A typical personalised page is shown in Figure 2.11. Here the system has recommended several music albums to the user. For each recommendation, the link "Why was I recommended this?" is provided. Clicking this opens another window, as in Figure 2.12, that displays similar products that were positively rated by the user, that caused the recommendation. That is, personalisation is explained in terms of the user's usage

history rather than their inferred interests or characteristics. However, the user has no access to how the usage history was used to provide the adaptive response, other than to assume there is some correlation between the products they have previously rated or purchased and the recommended product. The user can update their user model in several ways. They can specify they already own the recommended product or indicate their level of interest in it. They can also select which rated/purchased products may or may not be used to provide future recommendations. For example, if the user purchased a product as a gift for another person, they do not necessarily have any interest in it themselves so it might not be appropriate to use to generate recommendations.

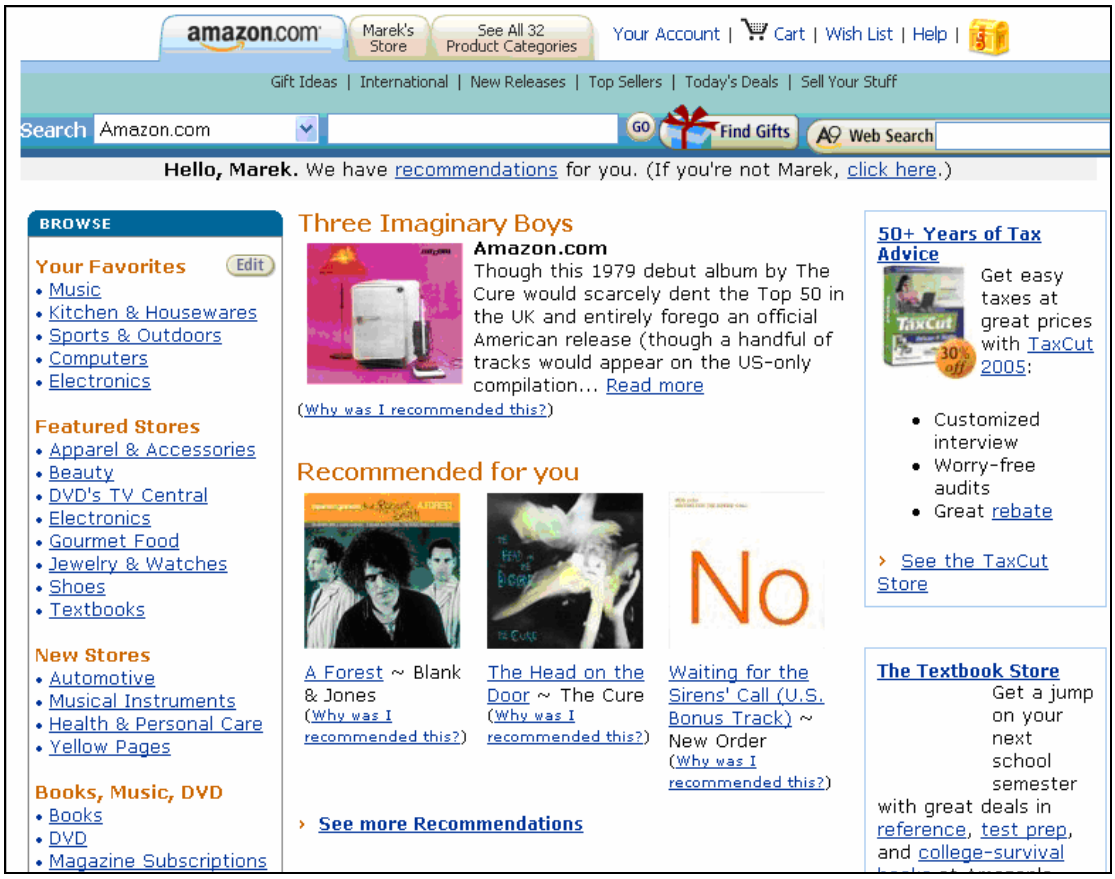


Figure 2.11 –Typical page on Amazon, showing product recommendations,

There are several key differences between this work and SASY. Firstly, as SASY is not a pure recommender system, the way it processes information about users to provide an adaptive response is different from recommender systems. SASY's personalisation is based on author-defined rules involving one or more user model attributes. SASY's adaptive content and rules are embedded tightly in to the page content. For example, a single paragraph may contain multiple adaptive words or links within it. The rules that govern when content is included/excluded may also be nested. The purpose of the personalisation is also different. Rather than providing different recommendations to different users, SASY's personalisation was designed to present the same information differently to different users. Thirdly, SASY allows the user to scrutinise content that was omitted by personalisation in

addition to content that was included. This allows the user to predict the impact of changing their user model in terms of the how it effects personalisation.

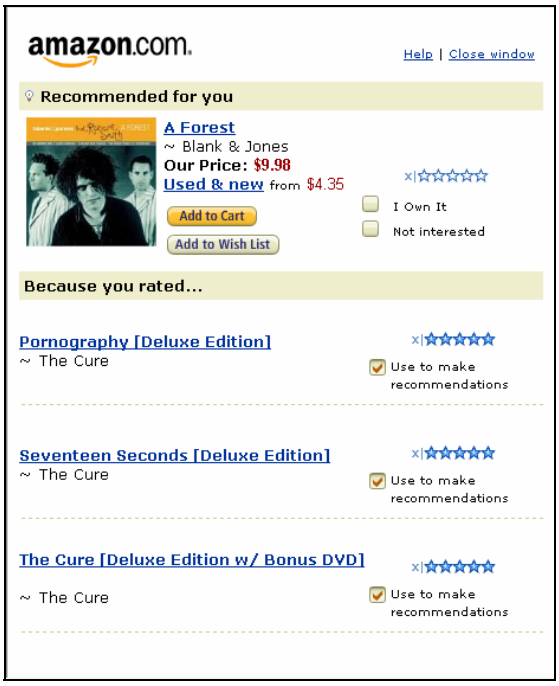


Figure 2.12 – Explanation page in Amazon, explaining why a product was recommended.

A commercial product that provides recommendation explanations similar to that of Amazon, is the Findory (http://findory.com)[11]. The Findory website is essentially a portal for news items and user blogs. As the user accesses news and blog items, the website infers interests about the user and makes recommendations for other news and blog items on the basis of this inference. The user is anonymous, but can scrutinise the system to obtain an explanation as to why an item was recommended. Clicking a special icon image next to a recommended item on a regular page (the sunburst icon following the text *Findory Recommended Article* in Figure 2.13), navigates to another page that lists the items accessed by the user that invoked the recommendation: the user can see and delete their history items accessed that influenced the recommendation. An example of an explanation page is shown in Figure 2.13. Here the user has requested an explanation as to why the article titled *Man shot dead inside SF business* was recommended to them. In the figure, the article was recommended because the user accessed the article titled *Shooting at church leaves woman dead, gunman killed*. A subtle point, yet important for this thesis, is that to access the explanation for a recommended item, the user to has to find and click on a special icon image, displayed next to the recommended item. The potential difficulty with this, based on evaluations of similar approaches with Tutor (see Chapter 5), is that the user might not be able to find this link or may notice it but not understand its purpose.

[11] The explanation facility of Findory was actually developed by the designer who developed Amazon's explanation facility (Greg Linden).

Figure 2.13 – Explanation page in Findory, explaining why a news item was recommended.

Tsandilas & Schraefel (2004) also studied a recommender system that allowed the user to scrutinise the personalisation. Their focus was on combining adaptive support with direct manipulation to make the system's adaptive behaviour more transparent, predictable and controllable. In their system, the user model captures the user's interests and indicates the level of each interest, for example, interest in dance or music. Content pages are segmented into fragments and each fragment has associated metadata that indicates what interests it relates to. Their system employs a fisheye adaptation technique, that essentially adjusts the font size of content so that it is proportional to the user's estimated interest in that content. An example of an adapted page is shown in Figure 2.14: the second to fourth paragraphs are shown in smaller font than the rest of the page because they relate to topics the user has little interest in, according to the system. The left hand side of the page displays the user's interests and their level of interest, using font size. For example, *Music* and *Film* are displayed in larger font than *Dance* indicating the user has more interest in these. Below that, is a slider element, called the *adaptation controller*, that is used to adjust the level of context: when the slider is at zero, adaptation has no effect on the appearance of pages, so all fragments are the shown in the same font size; when at maximum, non-relevant fragments are not shown on the page at all (as is the case in SASY), which means no context is provided. By default, the page does not show the interests that relate to each paragraph. However, the user can move their mouse over a paragraph to pop-up text that lists the relevant interests, for example, "Music, Dance". The user can also double click on an out-of-focus paragraph, this zooms in so the text is legible. Double-clicking again returns to the normal view. To control the adaptation, the user can adjust their interest level as follows: clicking an interest item in the left menu, for example, Music, pops up an another slider (see Figure 2.15) that allows the user to adjust

their level of interest. In summary, the fisheye adaptation technique allows the user to visualise and control the *focus* of the display (i.e. their interests) and see the *context* of the adaptation.

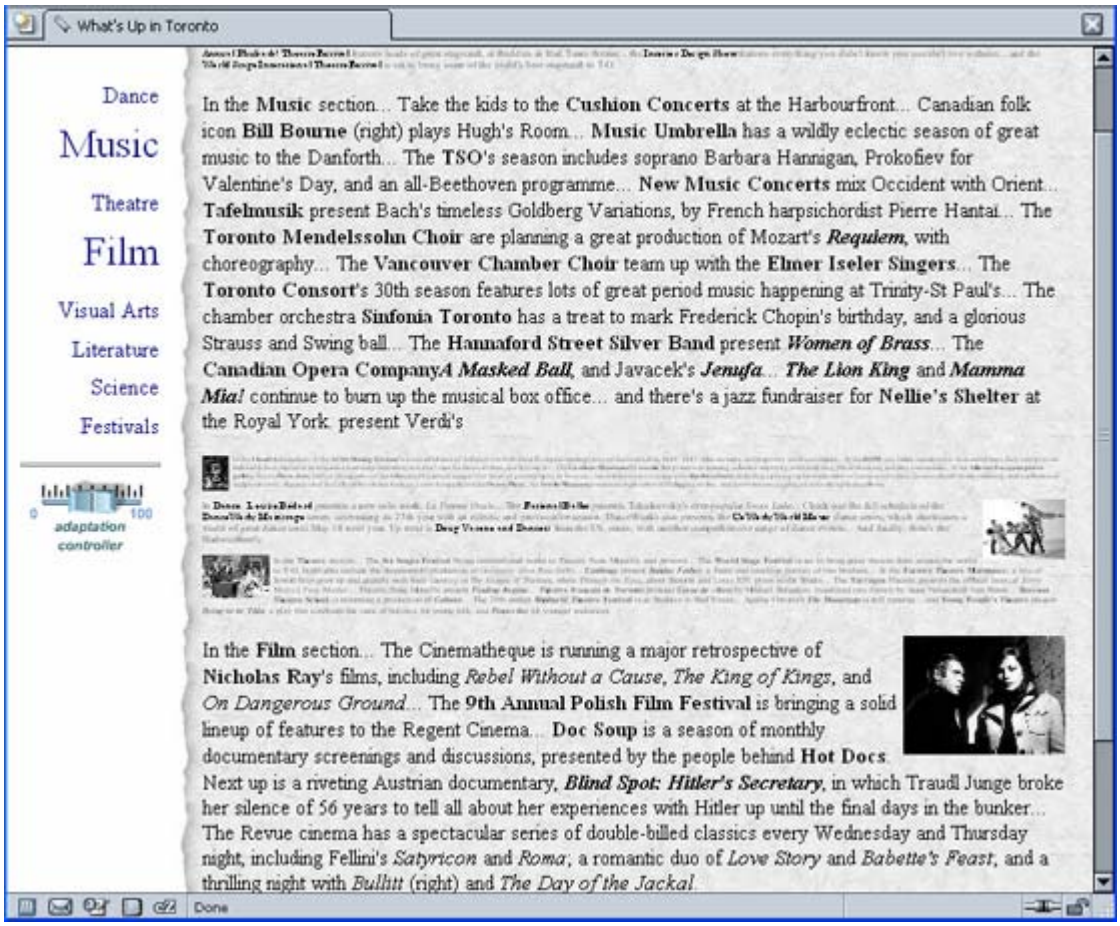


Figure 2.14 – Sample of an adapted page, using the fisheye technique as in Tsandilas & Schraefel (2004).

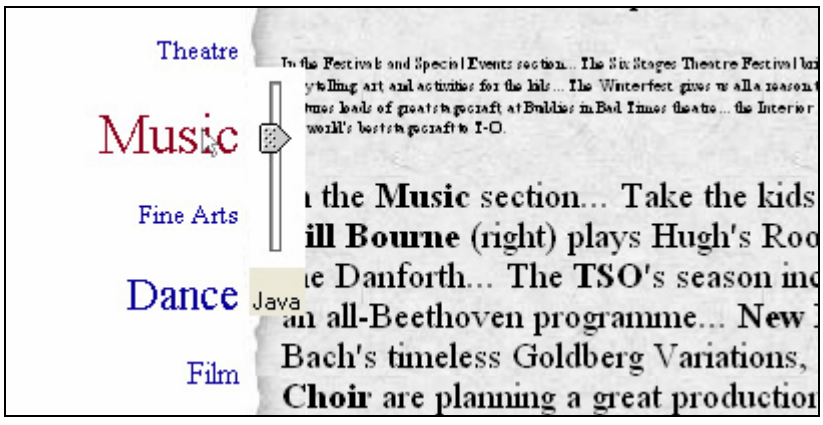


Figure 2.15 – Updating the user model in Tsandilas & Schraefel (2004).

The method for adapting content is different to SASY. Here, both the user model and the content of a page are described as a linear combination of vectors that represent topics of interest. The strength of this approach is that it could work well in systems where content is automatically classified by using

common text classification techniques. In other words, it would not require the author of an adaptive page to explicitly specify how the paragraphs within a page should be adapted: the adaptation is similar to that of a recommender system such that it models mainly user interests rather than arbitrary user model attributes, like SASY. SASY, on the other hand, provides the author in control of the adaptation, so the author specifies rules for the conditional display of content to the user, which can be more complex that just referring to interest levels.

In terms of scrutinisation tools, there are some similarities and differences with SASY. Both display user model attributes that affect personalisation as a main feature on the page: for example, Figure 2.14 lists the elements on the left of the page. In contrast, SASY lists the elements on the right and only lists those that effected personalisation on that page, so the list will vary from page to page. SASY uses text to describe the values of user model elements (e.g. You like Rock Music), while Tsandilas & Schraefel use font size to indicate interest levels. Tsandilas & Schraefel use pop-up text to describe the concepts relevant to text fragments on the page, SASY uses pop-ups to describe why items were removed or included in the personalised view of a page. Both systems allow the user to edit their user model: SASY requires the user to navigate to a user profile page, while Tsandilas & Schraefel allow the user to edit interest levels directly on the personalised page (see Figure 2.15). While both works have similar scrutinisation goals, each allows the user to scrutinise a different adaptation technique. In addition, SASY explains it adaptation rules to the user while Tsandilas & Schraefel aim to make the system's adaptation behaviour transparent without revealing the adaptation algorithm or mechanism.

2.5 Summary

This chapter has described related work in the fields of Adaptive Hypertext and Hypermedia (AH) and Open User Modelling, in the very specific areas that are relevant to this thesis. We have shown that the type of adaptive hypertext we aim to make scrutable in this thesis has been used in several research systems and current commercial systems: namely the inclusion and omission of adaptive content. There are few examples of adaptive hypertext systems that allow users to scrutinise other forms of personalisation, which suggest scrutability is important, however, not yet deeply researched. The scrutability in Amazon is arguably one of the most important since it is a commercial system.

The spirit of the scrutability made available in Amazon, IOS and Jameson's adaptive conference website is close to what we are researching in SASY. However, there are several elements that make our work novel, as defined in Chapter 1. Firstly, SASY allows the user to scrutinise the personalisation of an adaptive hypertext more deeply than what is reported in the literature. Secondly, the SASY is a generic, domain-independent framework for the creation of adaptive hypertexts. It is not domain or application specific. In addition, this thesis explores several ways to provide support for scrutinisation, qualitatively evaluates the scrutinisation tools developed and presents guidelines for developing such tools.

The next chapter presents the SASY scrutinisation tools in detail.

Chapter 3 SASY

This chapter describes SASY 4, the final system in the series of systems and system variants that have been developed to explore ways to support users in scrutinisation a personalised hypertext: Tutor 1, Tutor 2, Tutor 3, Cell-Tutor, SASY 1, SASY 2 and SASY 3. For a discussion of the evaluation of scrutinisation support tools, the reader is referred to Chapter 5. The system name SASY refers to SASY version 4, unless otherwise stated.

We first describes the type of adaptive hypertext this thesis aimed to make scrutable. Next we discuss the requirements and challenges involved in designing a Scrutable Adaptive Hypertext framework, and tools that allow an end user to scrutinise it.

The remaining sections describe SASY system architecture and the scrutinisation tools.

3.1 Requirements for the Underlying Adaptive Hypertext framework

The intention of this thesis is to explore scrutability in a typical and realistic adaptive hypertext. In other words, to explore how to support the scrutinisation types of personalisation that resemble what might occur in a real world personalisation system. This section describes the types of personalisation for which scrutability is explored in this thesis.

3.1.1 Overview of the Personalisation mechanism

The personalisation mechanism used is representative of a typical adaptive hypertext system (Brusilovsky, 2001). Essentially, the personalisation in SASY involves inserting/removing hypertext content fragments based on logical expressions involving values of user model attributes. This is a simple but widespread approach, so it makes a logical starting point for exploring ways to explore the scrutinisation of an adaptive hypertext.

The system builds a model of each user, populated initially by the user's answers to a brief questionnaire. The user model is later updated and refined by the system through inferences and observations made about the user.

To personalise a hypertext document, a user's user model is processed according to rules embedded in the requested document to determine what parts of the content should be added or removed to create the user's personalised view of the page. This is a standard Adaptive Hypertext technique (Brusilovsky, 2001), similar to the approach employed by AHA! (Calvi & Cristea, 2002) and was based on the approach presented by Kay & Kummerfeld (1994).

The rules for conditionally displaying adaptive content to a user may be simple logic expressions based on one or more user model attributes, or complex expressions developed in python code (Note: the syntax is discussed in Chapter 4 which presents ATML, the XML language in which content for SASY is written).

3.1.2 How the user model is maintained

A user's characteristics may change over time. For example, while reading a tutorial, the user will (hopefully) gain knowledge. Over a longer time period, a user's interests might evolve or change completely. A user model that changes, based on inferences or observations about the user, is a typical feature of an adaptive system (Kobsa, 2001), and this approach is used by this thesis.

The adaptive hypertext framework used in this thesis allows updates to be made to the user model based on the following user-triggered events:

- User access to a page;
- User mouse click on an HTML element;
- User selecting an answer to a quiz question.

As an example of how this might be used, consider a holiday recommendation system might attempt to generalise the user's interests based on holiday locations they click on. It might infer they like beach holidays or exotic locations and then recommend holiday deals to match these interests.

The rules for populating the initial user model and updating it based on inferences may be simple logic expressions based on one or more user model attributes or complex expressions developed in python code.

3.2 Requirements for a Scrutable Adaptive Hypertext

Given the adaptive hypertext framework described above, the following details the requirements for the scrutinisation components.

We define the *Scrutinisation Tools*, as components of the system interface that allow the user to inspect, understand and control the personalisation. The following are requirements for such tools.

Requirement 1. The scrutinisation tools must allow the user to scrutinise the personalised hypertext in the following ways:

- understand exactly what aspects of the hypertext were personalised;
- see what aspects of the user model caused each personalisation;
- view evidence of how user model attributes were set, whether set directly by the user or inferred by the system;

- understand and visualise the implications of changing their user model, in terms of how it affects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes;
- change aspects of their user model to control the personalisation.

Requirement 2. The scrutinisation tools must allow the user to understand and control both simple and more complex forms of personalisation, as defined below.

- Simple personalisation:
  - Personalisation based on the value of a single user model attribute.
  - Personalisation based on a user model attribute that was set explicitly by the user, such as a user preference.

- Complex personalisation:
  - Personalisation based on the result of complex expressions involving one or more user model attributes.
  - Personalisation based on a user model attribute(s) that have been inferred about the user.
  - Personalisation that changes over time based on a changing and evolving user model.

Requirement 3. The user interface should be easy to use without training to support users who may not use the tools often. However, it should also support "power" users who will use the tools often.

Requirement 4. Text explanations provided by the scrutinisation tools should be generated by the system where possible to reduce the effort in creating content. However, hypertext authors must have the ability to override default explanations if required, for example, to explain complex personalisation rules.

3.3 Key Challenges for Scrutinisation Tools

Walk-up-and-use Interface

One of the challenges of supporting scrutability is that we would expect users might only use this facility irregularly, based on our previous work with scrutinisation (Czarkowski & Kay 2000). Typically, users can be expected to focus on something else like retrieving information, learning, reading or doing what ever the adaptive system was designed for. Scrutinisation tools are not an element for which we would expect users to be trained as part of the normal use of the interface. Users might go for long periods without scrutinising. However, when there is a suitable trigger, such as unexpected behaviour of the personalisation, the user should be able to work out how to delve into the adaptation.

Supporting both Novice and Advanced users

Whilst most users may not scrutinise often, there may also be users who scrutinise very often. For example, Manber et. al. (2000) reported MyYahoo! Had a number of "power" users who had spent a great deal of effort in customising their site so that it suited their needs.

Visibility

Another key challenge, that relates to those mentioned above, is striking a balance for the obtrusiveness, or visibility, of the scrutinisation tools. One might expect if the tools will not be used often they should be hidden away. On the other hand, users must be able to find the tools when needed. If the tools are too unobtrusive, users may never find them when required.

Explaining personalisation to users

This thesis explores ways to explain the process of personalisation simply, so that it can be easily comprehended by end-users. One of the key questions is how much does the user need to be told to understand the personalisation that occurred and how to control it to better suit their needs.

Indeed, in order to understand personalisation, the user must also understand the following concepts:

- The system stores information about them.
- The information stored about them is based on information they volunteered but it is also based on inferences and system observations of their use of the adaptive hypertext and environment.
- The information stored about them is used to decide what content is presented to them.
- They are able to change the information stored about them and this changes the personalisation.

3.4 SASY Overview

SASY (Scrutable Adaptive hypertext System) is the last of several versions of scrutinisation tools developed as part of this thesis. It provides a framework for the creation and delivery of personalised content that the user can scrutinise, or inspect, to understand what personalisation occurred and how to control it.

To the end user, SASY is a Web application that presents personalised content that is accessed and scrutinised through a standard web browser. Authors create content by writing a set of XML documents that conform to a schema called Adaptive Tutorial Mark-up Language (ATML). An ATML document is an HTML 4.0 document with additional tags that allow the author to define personalisation rules for adaptive content, and is the subject of Chapter 4.

Content is grouped into topics, such as a UNIX Security tutorial or a Personalised television guide. For each topic accessed by a user, SASY builds a profile of the user's interests, background, goals and other personal characteristics from answers the user provides to an initial online questionnaire. The profile can be set up with defaults to reflect a stereotype. The profile may be updated by the system as user characteristics are inferred.

The mechanism by which a hypertext document is personalised works as follows. When a user requests a page, SASY evaluates personalisation rules embedded in the ATML page against their user profile. Thus, SASY determines what adaptive content should be added to the user's personalised view of the page and what should be omitted. The personalisation rules and what attributes to model about the user are determined by the author when the content is created. In this way, different users see a different view of the same content that has been adapted to suit their individual characteristics as defined in their user model (profile).

What makes SASY unique is its scrutinisation tools, that allow the user to see what personalisation occurred and how to control it. Each page includes a summary of the personalisation (e.g. number of content items added and removed). On request, the system will highlight every personalised item on the page and provide an explanation of why each personalisation occurred. The system also describes how and why each profile attribute was set. This makes the system accountable for profile attributes that have been inferred about the user based on their actions.

SASY and ATML evolved from a series of systems (described in Chapter 5) that were initially used for the presentation of adaptive teaching material. Although SASY and ATML are domain independent, both include features that are particularly useful for online learning. For example, ATML includes tags that allow the author to easily define quiz questions. The design of the SASY interface was strongly inspired by ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001), an application that teaches the computer programming language LISP.

This chapter provides an overview of the SASY framework and describes the end user's view and the scrutinisation tools in detail. The following sections present the system architecture, the user's view of SASY, the user model and scrutinisation tools.

3.5 SASY System Architecture

The SASY System Architecture is illustrated in Figure 3.1. Users access the SASY web application through a web browser. SASY 4 supports the following the following web browsers that support JavaScript 1.2 and cookies: Microsoft Internet Explorer 6+, Netscape Communicator 7, Firefox 1. The SASY server is implemented as a collection of Python CGI (Common Gateway Interface) scripts. This platform was selected because it allows rapid development. The scripts implement the following components: Presentation Server, Personalisation Server and the User Model server.

Each page request is handled by the Presentation Server. This manager component is responsible for dynamically building an HTML document for the requested content embedded within the standard SASY user interface mark-up. It first retrieves and parses the ATML document requested. It also retrieves the user model for the current user through the User Model server.

Where the ATML document contains adaptive content, the Personalisation Server component is invoked to determine whether the content should be added or removed from the rendition of the page that is shown to the user. The Personalisation Server evaluates the expression contained in the ATML adaptive mark-up against the user model, through the User Model server. The result determines whether the content is added or removed and at this time, the Personalisation Server generates a textual explanation that will be shown to the user to explain the personalisation that occurred.

The ATML document may also contain commands to retrieve or update values in the user model. The User Model server is invoked to handle these requests. A separate user model is maintained for each topic accessed by a user. This is because the design assumes each SASY topic may be developed by a different author, may generally contain different attributes or have different meanings for the same attributes. Each user model is stored as a binary file.

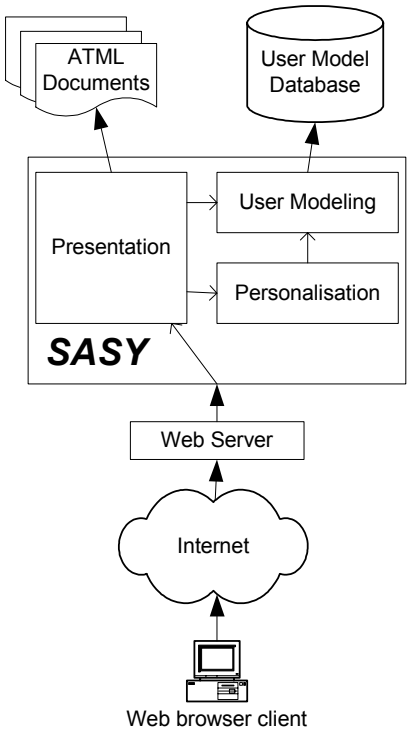


Figure 3.1 – SASY System Architecture

3.6 Initial Profile Creation

Upon logging in to SASY, the users selects from the list of available SASY topics. On entering a topic, SASY presents a welcome message explaining that SASY will personalise content, and requires the user to complete a questionnaire form to populate the initial user model.

Figure 3.2 shows the profile page from a Personalised Holiday planning topic in SASY. This topic provides holiday recommendations based on the user's interests and preferences as determined from their profile. Here, to build the initial profile the user is asked whether they are *Single*, a *Couple* or *Have Kids*. It also mentions attributes that will be inferred by SASY later, for example, *Additional holiday preferences*, but states there is *Nothing observed yet*. This reminds the user that SASY may infer things about them.

The user's answers are used to build their initial user model for the topic. The model also contains attributes that are calculated based on arbitrary logic expressions involving other user model attributes. For example, the author may define a rule to determine the user's budget as follows: assume the user has a high holiday budget if they are *Single*, assume *Couples* have a lower budget per person.

Users who have previously accessed the topic have the option of updating their previous answers, as well as changing attributes that have been inferred by SASY.

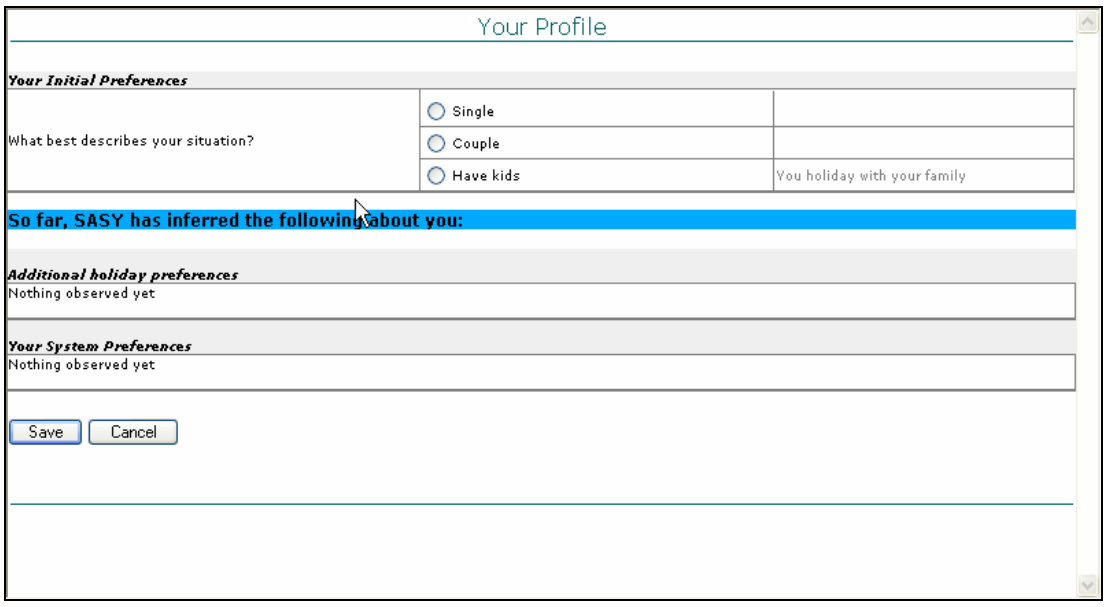


Figure 3.2 – Typical Profile page in SASY.

3.7 Content Pages

On saving the profile, SASY navigates to the Home page for the topic. This is a typical content page but it is the navigational hub of the topic because users will always start at this page at each new session in the topic.

An example of a typical content page is shown in Figure 3.3. Content pages all contain the same menu bar at the top of the page. Starting from the left of the page, the menu has across the top:

- Home – opens the Home page.
- Contents – opens the Contents page that shows an adaptive list of all pages in the current topic.
- Your Profile – opens the Profile page.
- Make Notes – opens another window where the user can make personal notes that are stored by the system.
- Change Topic – opens the topic list so the user can access a different topic.
- Help – opens the online help page.

Figures Figure 3.3 and Figure 3.4 illustrate how the same content page might be presented differently to different users. The page in Figure 3.3 provides a holiday recommendation personalised to a user who in a relationship (i.e. *couple*) and has a *low* holiday budget. Notice how the first paragraph contains *This is a great opportunity to spend some romantic, quality time with your partner*. By contrast, the page in Figure 3.4 does not have this text because it has been personalised to a user who is *single* and has a *high* budget. Also, the second paragraph in Figure 3.4 is different because it presents accommodation that has been tailored to the a higher budget.

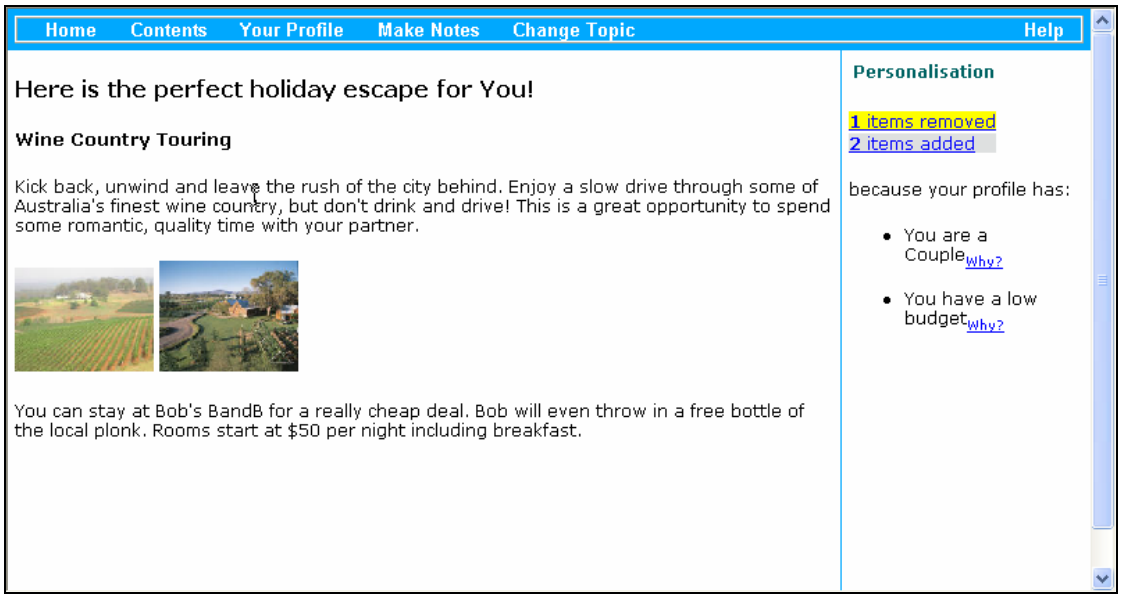


Figure 3.3 – Example of a typical content page. The page shows a holiday recommendation for a *couple* with a *low budget*.

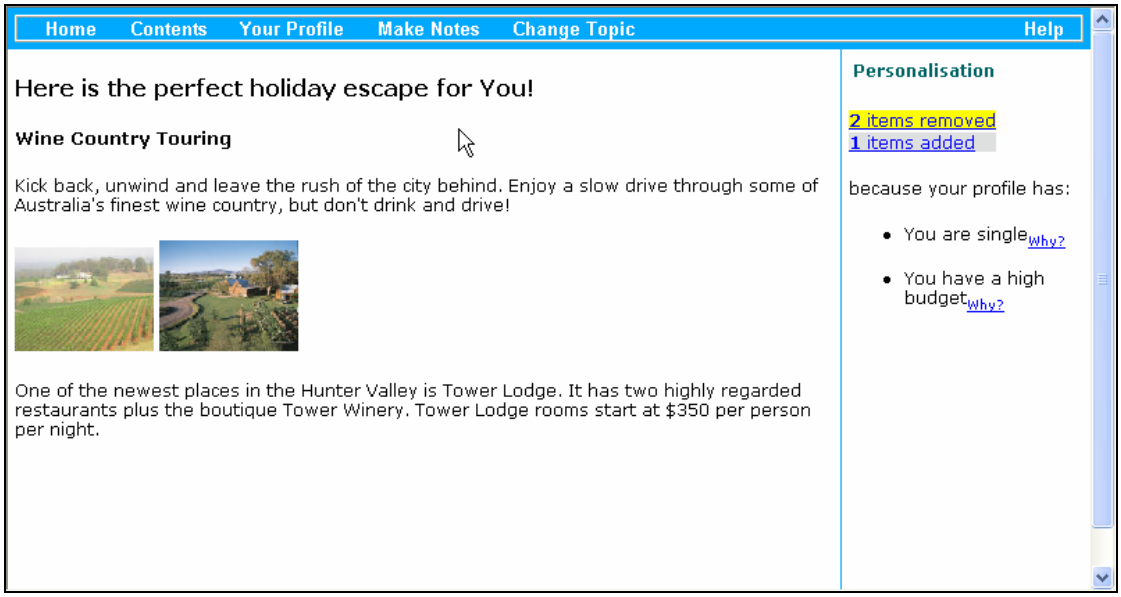


Figure 3.4 – Example of a typical content page. The page shows a holiday recommendation for a *single* user that has a *high budget*.

Content pages may contain a mixture of static and adaptive content. The adaptation performed by SASY can be coarse-grained or fine-grained. Adaptive content can be of arbitrary size, be it a whole page, paragraph, word, or even a single HTML element. This gives the author flexibility as they can write adaptive words mid-sentence rather than providing two versions of the paragraph which would

otherwise be the same. Since HTML browsers support different forms of media (text, sound, image), it is straightforward to adaptively include or exclude these forms of media based on the user's profile.

Content is marked as adaptive by the author, who specifies a logical condition in the ATML mark-up. This condition must match a user's profile in order for that user to be shown the content. For example, in the holiday recommendation above (Figure 3.4), the author specified that the paragraph describing the expensive *Tower Lodge* accommodation should only be shown to users who have a high holiday budget. This is a simple condition. However, complex logic conditions may also be specified through python code expressions involving other user model attributes.

A user's characteristics may change over time. For example, while reading a tutorial, the user may gain knowledge. Over a longer time period a user's interests might evolve or change. In SASY, the content author can specify rules (through python code expressions embedded in the ATML) for updates to the user model based on the following user-triggered events:

- User access to a page.
- User mouse click on an HTML element.
- User selecting an answer to a quiz question.

For example, the holiday recommendation topic might attempt to generalise the user's interests based on holiday locations they click on to read about. It might infer they like beach holidays or exotic locations and then recommend holiday deals to suit their interests.

While personalisation allows the presentation of content that is tailored to the individual, it also introduces the need to allow the user to scrutinise the adaptation. The user might want to find answers to questions like:

- Why does it think I like the beach?
- Why does it think I have a high budget?
- What accommodation would it recommend if I had a low budget?
- How can I tell it I hate the beach?

SASY supports the users to find answers to such questions through its scrutinisation tools:

- **Highlight tool –** allows the user to see what parts of a page have been personalised, understand why the personalisation occurred and how to change it.

- **Evidence tool –** provides explanations about how/why/when user model attributes were set.

- **Profile tool –** allows the user to view and edit aspects of their profile.

The next sections describe these tools in detail.

3.8 Highlight Tool

The right hand side of every content page (e.g. Figure 3.3 and Figure 3.4) is called the Personalisation panel. The scrutinisation tools are invoked by clicking elements in this panel. In Figure 3.3, the panel contains the hyperlinks: *1 items removed* and *2 items added*. This specifies how many content items have been removed and added by the personalisation on that page, providing a summary of the personalisation that occurred on a page.

To see what was personalised, the user clicks these links to invoke the Highlight tool. This reloads the current page to show all the content available for that page, annotating the content that was added and removed to create the personalised view of the page. It applies a *yellow* background colour to content items that were removed by personalisation and a *grey* background colour to content that was added to the personalised view. This colour code is described in the key at the top of the page. For example, if the user invokes the Highlight tool from Figure 3.3 the page reloads and appears as in Figure 3.5. Notice how the adaptive text *This is a great opportunity to spend some romantic, quality time with your partner* in the first paragraph is given a grey background colour, indicating it was added by personalisation.

The Highlight tool also explains why each personalisation was applied. The first grey section on the page includes the explanation text *Added because profile has: You are a Couple*. The means content highlighted in the grey area was added to the personalised view because the user's profile states they are a couple. Likewise, the content in the second grey area was added because the user has a low budget. The yellow area highlights content that was removed from the personalised view because the user has a low budget, according to their profile.

The explanation text for each personalisation informs the user which profile attribute values from the profile caused the personalisation. The user can always change what is added and removed by personalisation by changing their profile.

Figure 3.5 – Example of the Highlight tool in default mode.

What is described above is the default configuration of the Highlight tool in SASY 4. In the default mode the explanation text, e.g. *Added because profile has: You are a Couple*, is displayed on the page itself, just above the actual content that was added or removed. There is an another mode of the Highlight tool, called the Mouse-Over mode. In this mode, added and removed items are still annotated with grey and yellow backgrounds, but the explanation text is not included on the page and is initially hidden. Figure 3.6 shows an example of the Highlight tool in mouse-over mode. Notice the instructions at the top of the page directing the user to use move their mouse pointer over highlighted items.

If the user moves their mouse over a highlighted area, the explanation text pops up over the page, then disappears again when the mouse was moved away. For example, Figure 3.7 shows the pop-up text "Added because your profile has: You are a Couple". Earlier versions of SASY (1 through 3) only featured the Mouse-Over mode.

SASY 4 allows the user to access either the default or mouse-over form of the Highlight tool. This is configured through a setting in the profile. The advantage of the mouse-over mode is that it shows what

personalisation occurred through colour annotations without cluttering the screen with explanation text. This makes it feasible for the user to read through every page with the highlighting on. The advantage of the default mode is that is provides full detail of why each personalisation occurred without requiring the user to further interact with the page using their mouse.

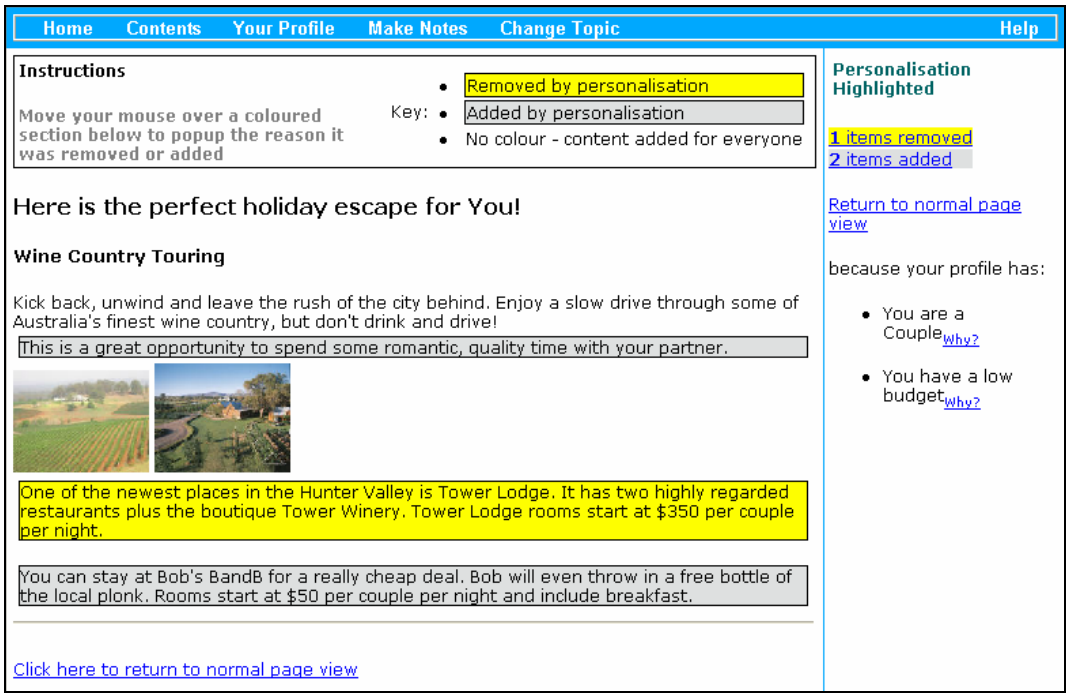


Figure 3.6 – Example of the Highlight tool in mouse-over mode.

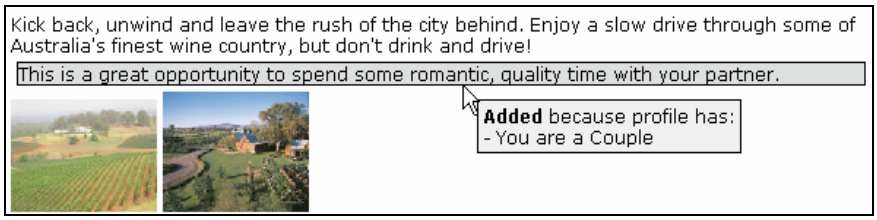


Figure 3.7 – Example of the pop-up explanation text of Highlight tool in mouse-over mode.

3.9 Evidence Tool

When personalisation occurs on a page, the right hand side of the page lists attribute values from the profile that were used to make decisions about what to add and remove from the page. The list will generally vary from page to page, as it is likely different aspects of the profile will drive personalisation on different pages. For example, Figure 3.5 lists:

- You are a couple
- You have a low budget

Each element in the list is followed with a hyperlink labelled *Why?* Clicking this pops up the Evidence tool, that explains how, why and when the attribute was set in the profile. For example, when the user clicks for the *Why?* link for evidence about the belief *You have a low budget* (from Figure 3.3 or Figure 3.6), SASY opens the Evidence tool as in Figure 3.8. The Evidence tool shows *SASY assumed you have a low holiday budget because you are a Couple and are probably saving for a house or paying off a mortgage*. In this case, the author defined a rule in the initial profile creation to assume the user has a low budget if they select they are a Couple in their initial profile. This rule makes quite a bold assumption and we do not suggest it is fair. However, it is the sort of assumption that might be made, an over-generalisation by the adaptive hypertext author. This is one of the potential sources of the need for SASY style scrutability.

The Evidence tool includes a link to the profile and reminds the user they can control personalisation by changing their profile.

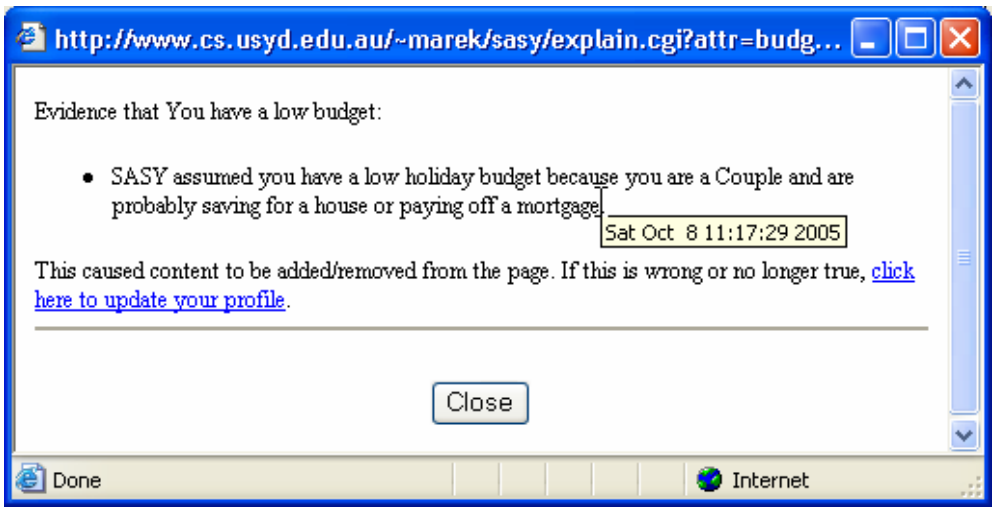


Figure 3.8 – Example of the Evidence tool.

3.10 Profile Tool

Earlier we introduced the profile page and showed how it appears when the user starts a SASY session (Figure 3.2) to populate the initial profile. The profile can later be accessed from the Evidence tool, or by clicking the "Your Profile" link in the menu. When accessed following the initial profile creation, it shows all the initial profile attributes as well as any attributes that have been inferred or observed about the user. An example is shown in Figure 3.9. The first section of the page shows the initial preferences that were set by the user. Underneath the heading "So far, SASY has inferred the following about you", it contains attributes that have been inferred about the user.

The profile page has three columns:

- The left column presents the profile question or attribute description. For example, the third row in Figure 3.9 has *What is your holiday budget?*
- The middle column has the available options for that attribute. For example, the third row in Figure 3.9 has *Low budget* and *High budget*. The current value is selected and the label is a hyperlink. Clicking the link invokes the Evidence tool to explain how the value was set. For example, Figure 3.8 is displayed when the user clicks *Low budget* from Figure 3.9.
- The right column has additional information that explains what the attribute value means, how it is used or how it affects personalisation in general. This captures the adaptive hypertext author's reasoning about the value shown in the middle column.

The profile page was designed to support users who scrutinise primarily through the profile. For example, the user might scrutinise by accessing the profile by clicking Your Profile in the menu of a content page. The user might ask *What is the effect of changing my profile? Why does an attribute have its current value?* The intent was to make the profile self explanatory and self describing. To achieve this, in part, the right hand column describes, at a very high level, the effect of each value on the personalisation. For example, next to the *Low budget* value it states "*You want a quality holiday but are sensitive to the cost. You will not be shown holidays dearer than $5000*", while for *High budget* it has *"You will be shown all holidays regardless of cost"*. These descriptions generalise the effect of selecting the corresponding value for the profile attribute on the personalisation.

The profile also contains several standard system preferences that allow the user to control how the scrutinisation tools function, how intrusive they are and how much help is provided to users when accessing the tools. These are shown as the last four rows in Figure 3.9, under the heading Your System Preferences.

- The first option controls whether the Highlight tool includes explanations on the page (as the default) or whether to use pop up explanations on a mouse over event (mouse-over mode).

- The second controls whether the list of profile attributes that affected personalisation should be shown on the page at all (as described in the Evidence Tool section). This allows users who are not interested in scrutinisation to hide this detail from their view altogether.

- The third allows the user to turn on the Highlight tool by default on every page. This shows all the available content with no personalisation applied but still annotates the adaptive content by highlighting what would be added or removed by personalisation. The annotation provides guidance while allowing the user to see the non-personalised version of every page.

- The fourth system preference controls whether the system shows additional information for using the highlight tool. If turned on, the system pops up a help window when the user invokes the highlight tool. Users who do not often use the scrutinisation tools might find this feature helpful.

3.11 Contents Page

The Contents page is accessed by clicking "Contents" in the top menu of any page. It is typically used as an adaptive site map of the SASY topic. It can also contain adaptive content in the same manner as a typical content page.

Hyperlinks in SASY use colour coding to provide additional navigation support. As in ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001), colour coding of hyperlinks is based on a traffic light metaphor. For any hyperlink, the author may specify zero or more pre-requisite pages. If the user has visited all the pre-requisite pages, the hyperlink will appear in green, indicating the user is ready to read the material. If the user has not visited all the pre-requisite pages, the hyperlink appears in red. If the user has already visited a page, its hyperlink will appear in black, but still underlined to distinguish it from normal text. Hyperlink colour coding indicates which pages the hypertext author considers the user should read or not, based on the current state of the user's user model. At the same time, the user is free to access any page.

A portion of the Content page is shown in Figure 3.10.

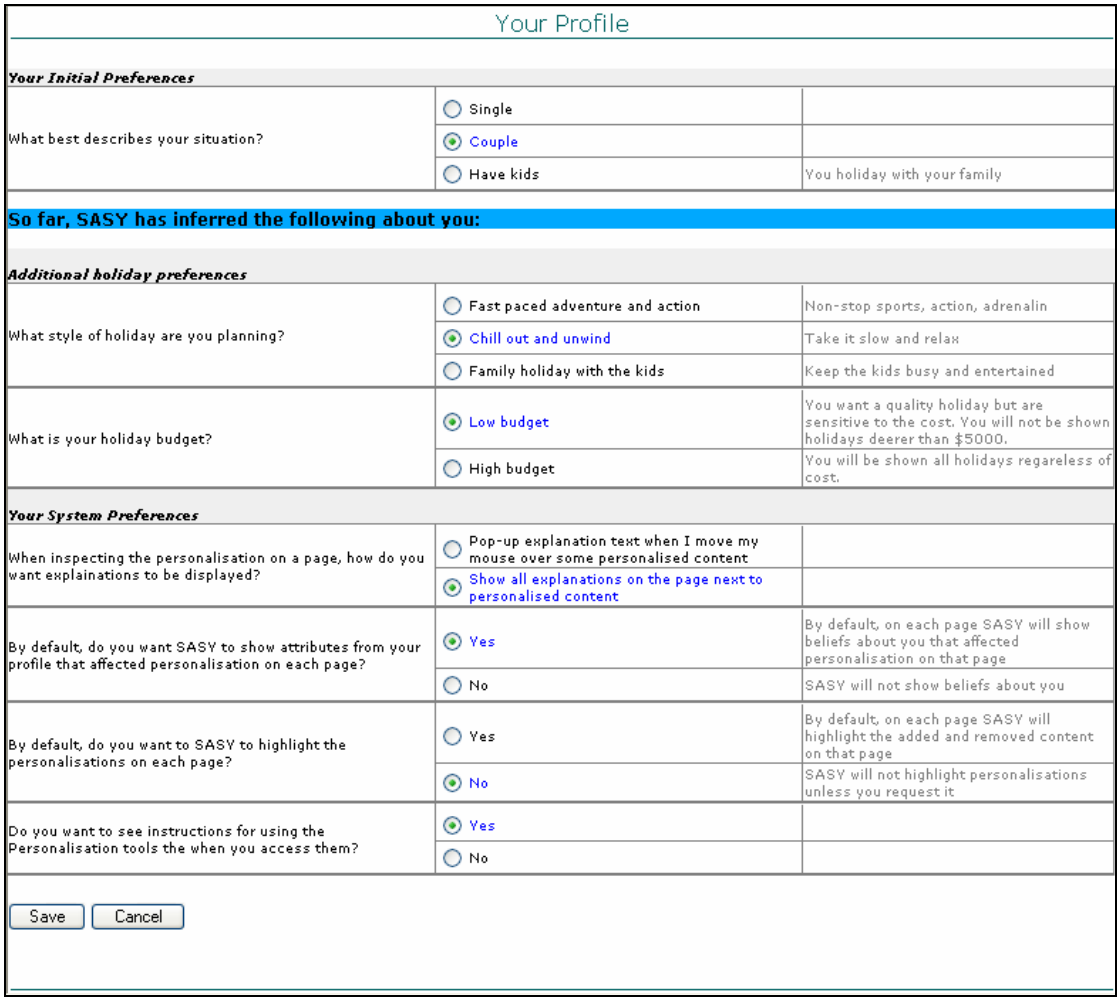


Figure 3.9 – Example of profile page on subsequent access.

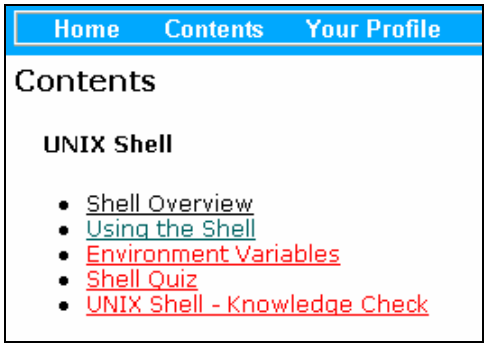


Figure 3.10 – Example of hyperlink annotation (see note below).

[**Figure 3.10 Notes**: Here the user has already accessed the *Shell Overview* page as this link appears in black font. The user has not yet accessed *Using the Shell*, but is ready to do so because it appears in green. All other links, for example *Environment Variables*, are pages the user is not ready to read and they appear in red. Perhaps the system believes the user has not yet learnt concepts that are pre-requisites for reading these pages. Here the annotation is guiding the user to next read the page *Using the Shell*.]

3.12 Notes editor

The Notes editor enables the user to create and store their own notes while they are working through a topic. Each time the notes editor is accessed the topic name and current timestamp are appended to the current notes. The notes are stored by SASY and displayed again when the notes editor is accessed. An example of the notes editor is shown in Figure 3.11. This page may be useful for summarising key points to remember, or things that seemed confusing and need follow up. It is available as a separate window for use in conjunction with navigation through content pages.

Clicking the *Save and Print* button parses the notes (which can be marked-up in HTML) and renders a printable page that the user can print or save to their local machine. The rendered version of Figure 3.11 is shown in Figure 3.12.

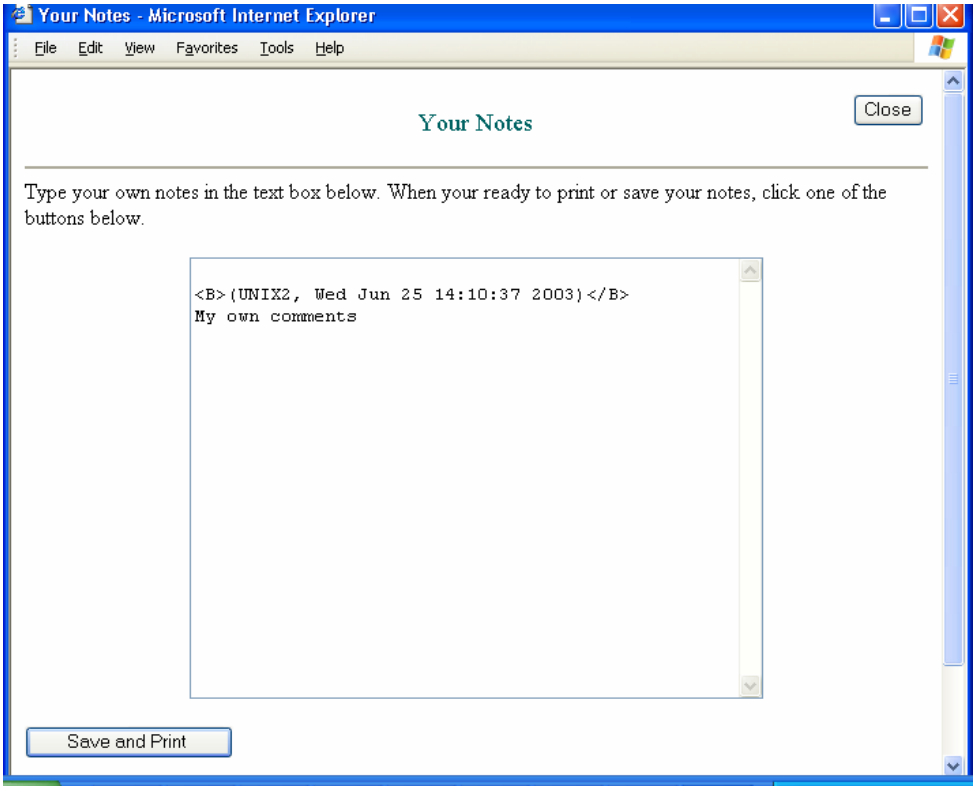


Figure 3.11 – Notes editor showing notes in raw form as entered by the user.

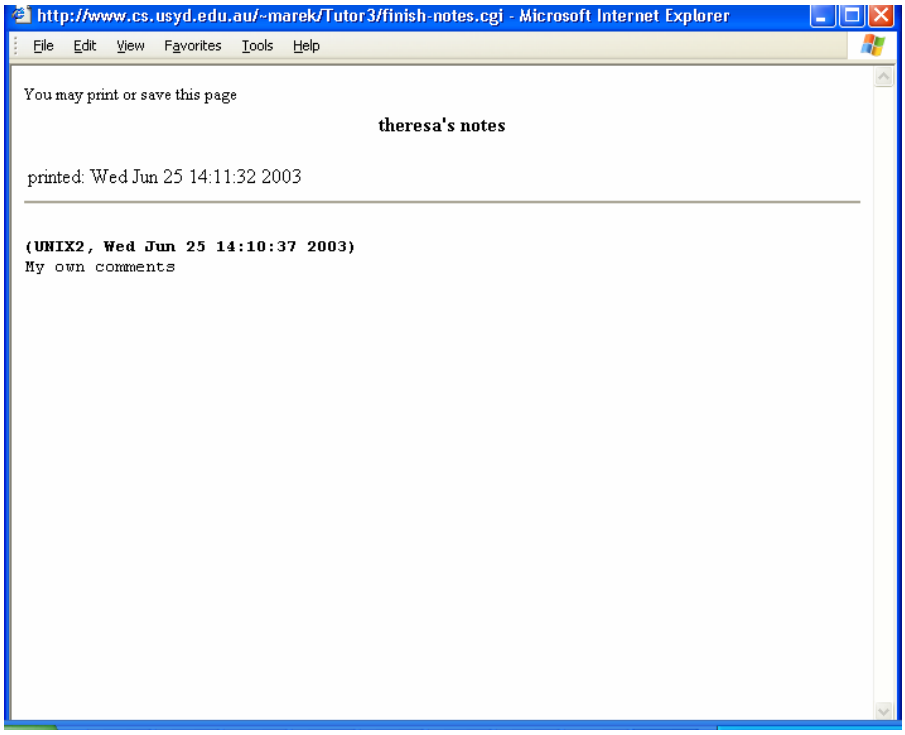


Figure 3.12 – Example of raw notes in Figure 3.11 rendered into a printable version

3.13 Summary

SASY's scrutinisation tools allow the user to find answers to questions they might have about the personalisation, be this general, or specific to the content domain: for example, for the holiday recommendation the user can gain answers to questions like:

- What has been adapted to me?
- Which holidays did it recommend based on the assumption that I like the beach?
- Why does it think I like the beach?
- What other holidays might it have shown me, but it did not because it treated them as irrelevant to me?
- What holidays did it show to someone else?
- How can I control or alter the adaptation?

This chapter provided an overview of SASY from the end-user's perspective. The next chapter describes ATML, the XML language used to create content for SASY.

Chapter 4 ATML

This chapter describes the Adaptive Tutorial Mark-up Language (ATML). ATML defines the structure and syntax rules for SASY content.

We begin with an overview of ATML version 2, which is used by SASY4. Then the ATML structure of the Profile page and Content pages is presented in detail.

4.1 ATML Overview

ATML allows the author to specify the metadata that is used by SASY to personalise content to the user and provide input to the SASY scrutinisation tools:

Personalisation Metadata – defines rules about how to personalise content to users and rules for updating the user model based on user actions. The approach is similar to that of the AHA system (Calvi & Cristea, 2002).

Scrutinisation Metadata – data used by SASY's scrutinisation tools that allow the user to understand and control the personalisation that has taken place.

This approach abstracts the complexity of the personalisation engine and scrutinisation tools away from the content author. The author only specifies metadata; SASY uses the metadata to generate a personalised HTML document and the JavaScript code necessary for the SASY user interface and scrutinisation tools.

ATML is an XML 1.0 document format and supports HTML 4.0 content. This facilitates easy migration from existing non-adaptive HTML content to a form that is both adaptive and scrutable.

Adaptive Tutorial Mark-up Language (ATML), as the name suggests, was first used for the adaptive presentation of teaching material. Most of the features and constructs are generic and not only applicable for authoring teaching content but other types of content as well. However, it still includes features that are particularly useful for the presentation of teaching material, for example the *question* tag that is used to generate multiple choice questions.

For large-scale use of SASY and ATML, it is envisaged that an authoring tool would be helpful to simplify the creation of adaptive content. Authoring tools have been created by researchers for similar forms of adaptive hypertext (De Bra et.al. 2003; Cristea, Aroyo 2002). It is feasible to create ATML in a text editor although this requires a technical competence. The ATML used by the evaluation of SASY (Chapter 6) was created this in this way.

4.2 SASY Topics

SASY content is segregated into topics, such as a UNIX Security tutorial or a Personalised television guide. Each topic requires a separate set of documents that define content for that topic.

To create a new SASY topic, the author creates the following ATML documents and publishes them to the ATML file system store:

- Profile page (um.xml) – defines the user model for the topic.

- Contents page (coursemap.xml) – defines an index page providing hyperlinks to content pages within a topic. The links are colour coded to provide of adaptive navigation support. The page may contain adaptive content so that different users see a different contents page.

- Home page (teacher.xml) – defines a base access point for the topic. Each time a user logs in to the SASY topic this is the first page they are shown.

- Content page(s) (*pagename*.xml) – defines a content page. Each topic contains zero or more content pages. The content may include adaptive mark-up, so that the same page is presented differently to different users based on their user model.

Each topic also requires the following static HTML pages that form part of the infrastructure of the user interface for the topic:

- index.html – defines a frameset layout for the interface.
- control.html – defines additional JavaScript functions required by the interface.
- base.css – a style sheet definition document (CSS1) where the author may specify their own style classes which can be applied to HTML tags used in the ATML documents.

4.3 Embedded HTML content

ATML documents may contain HTML 4.0 content. However, HTML tags used must conform to XML 1.0 schema standards. In particular:

- A close tag must be provided for every open tag. For example, <P> and </P>.
- Empty tags must closed with a "/", for example <IMG src="file.jpg" />.
- Attributes must be enclosed by quotes, as in the above example.

The remainder of this chapter presents ATML in detail. For the sake of brevity, examples are presented only at the end of each section to demonstrate the use of the key ATML tags.

4.4 Profile Page

We first present the tag structure of the profile page, briefly describe the key tags and illustrate with an example. Details for each tag are provided in the appendix 8.1.1.

4.4.1 Document Structure

The profile page document (um.xml) defines all the user model attributes that are stored about a user for the topic. Each attribute is defined using the *attr* tag. The profile page allows the author to define an initial questionnaire to directly elicit certain user characteristics (*initial* tag), placeholders for user model attributes that will be inferred later (*inferred* tag) and rules for defaulting user model attributes to build a stereotype of the user (*init* tag).

Figure 4.1 shows the tag structure of the profile page. The page is divided into three parts. The first part defines the course name (*course* tag). The second part defines user model attributes (*attr* tags) that are elicited from the initial profile (*initial* tag) or inferred later (*inferred* tag). The third part defines rules for initialising inferred user model attributes (*init* tag).

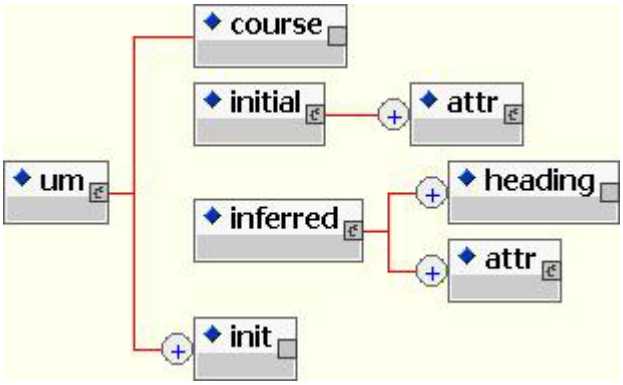


Figure 4.1 – Profile page element structure

4.4.2 Attr tag

The *attr* tag defines the metadata for a user mode attribute, including the data type, how the value is elicited, the range of values it may take and information that is provided to the user to allow them to scrutinise meaning of the attribute on the profile page and on content pages.

Attribute Data Types

The following user model attribute data types are supported by SASY:

- Fixed values. For example, “Yes” and “No”.
- Probability. An integer (0-100) that represents the probability that a belief about the user is true.
- Quiz Score. A counter, starting at zero, incremented if the user answers a quiz question correctly (refer to the *question* tag for details about quiz questions).

Fixed value and probability attributes can be set by rules on the profile page (see *init* tag), or inferred and updated through rules embedded in content pages (see *updateum* tag). Quiz Score attributes are only set and updated when the user answers quiz questions on content pages (see *question* tag).

Attr tag structure

The *attr* tag structure is shown in Figure 4.2. The *id* attribute is the internal name for the attribute that is used to refer to this attribute in other tags (e.g. *adapt* tag on content page) and rules (e.g. *init* tag on profile page).

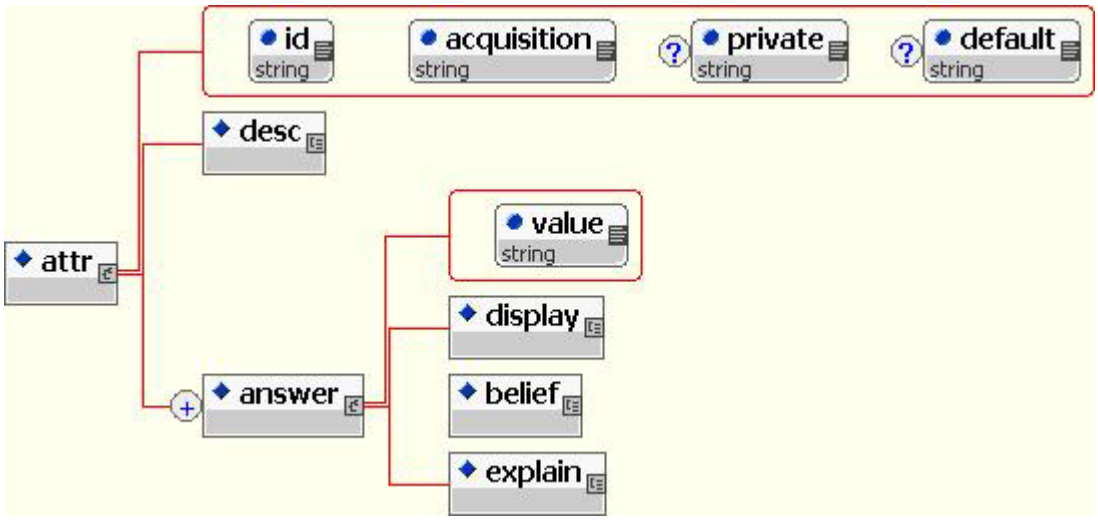


Figure 4.2 – Profile page – attr element structure

The sub-element *desc* tag specifies a description of the attribute that is displayed in the first column of the profile page in SASY. For example, "What is your holiday budget?" in the first column in Figure 3.9.

If the attribute is not a quiz score or probability type, a set of fixed values must also be specified. The *answer* tag is used to define each value that may be assigned to the attribute and has the following elements:

- *value* attribute – defines the value stored internally for the attribute. For example, "low" or "high". This is not shown to the user but is used to refer to the value in rule expressions in the *init*, *adapt* and *updateum* tags.
- *display* tag – this is the text displayed to the user to describe the value, shown in the second column on the profile page in SASY. For example, the second column in Figure 3.9 (pg. 90) has the display values "Low budget" and "High budget".

- *belief* tag – used to provide a brief description of the value, shown on the right hand panel of a content page in SASY, if the attribute affected personalisation on that page. For example, the right hand column of the page in Figure 3.3 (pg. 82) shows the belief value "You have a low budget".
- *explain* tag – defines an explanation of the value that is shown in the third column on the profile page in SASY. This should be used to provide text that explains how the *answer* will effect personalisation. For example, for the value "Low budget" in Figure 3.9 (pg. 90), the third column has the text "You want a quality holiday but are sensitive to the cost. You will not be shown holidays dearer than $5000".

4.4.3 Init tag

The *init* tag is used to define a rule that specifies how an inferred user model attribute should be defaulted or updated when the profile page is saved by the user in SASY. This tag is primarily used to specify defaults for inferred user model attributes, but can also be used to changes attributes if the user changes their profile because the rules are evaluated every time the profile is saved.

The rule can be configured to fire always when the profile page is saved (by setting attribute cond="default"), or only when a certain condition is true (by setting attribute cond=*some expression*). In both cases, the user model attribute is only changed if the user has not explicitly set the value for the attribute on the profile page. This allows the user to override defaults.

The cond attribute may be specified as an expression involving one or more user model attributes through Python-style code expressions. For example, the condition *self.budget=='low'* evaluates to true if the user model attribute *budget* has the value *low*. The prefix *self* is a Python reserved word that must be used when referring to a user model attribute.

Expressions may also use logical operators: *and*, *or*, *not*. For example, the condition *(self.budget=='low')and(self.status=='single')* evaluates to true if the user model attribute *budget* has the value *low* and the user model attribute *status* has the value *single*.

4.4.4 Profile page Example

We now illustrate the ATML for the profile page with an example, that describes the mark-up used to generate the initial profile page in Figure 4.3 and full profile page in Figure 4.4. This example defines a profile page for a Personalised Holiday planning topic in SASY called "HOLIDAY". This profile page is similar to the examples in Chapter 3 but has been simplified slightly for the purposes of this description. In this example, the rules used for creating a user profile make some over-optimistic model assumptions, but this is done so to illustrate the features of ATML.

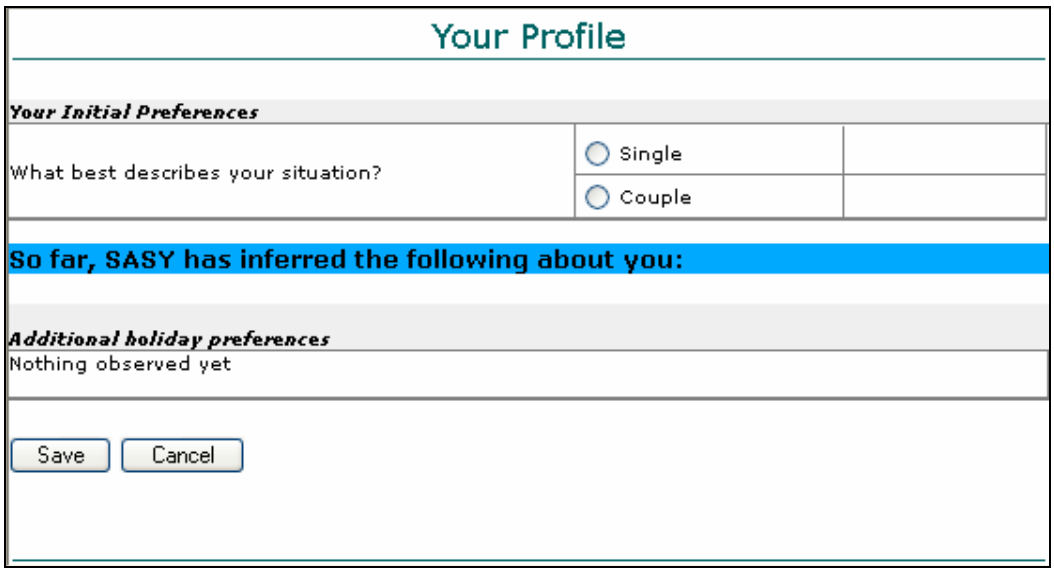


Figure 4.3 – Initial profile page example.

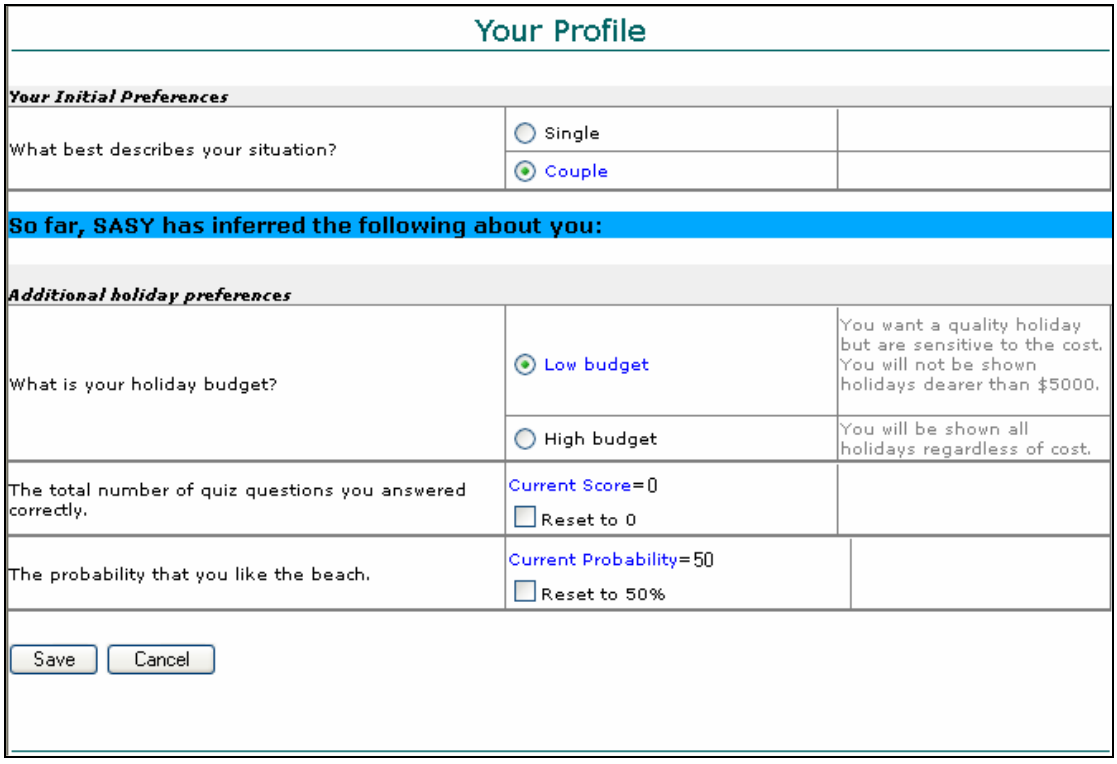


Figure 4.4 – Typical profile page once initial profile has been created.

```
1  <?xml version="1.0"?>
2  <um>
3  <course id="HOLIDAY" />
4  <initial>
5
6  <attr id="group" acquisition="direct">
7  <desc>What best describes your situation?</desc>
8  <answer value="single">
9       <display>Single</display>
10      <belief>You are single</belief>
11      <explain></explain>
12 </answer>
13 <answer value="couple">
14      <display>Couple</display>
15      <belief>You are a Couple</belief>
16      <explain></explain>
17 </answer>
18 </attr>
19
20 </initial>
21
22 <inferred>
23
24 <heading text="Additional holiday preferences"/>
25
26 <attr id="budget" acquisition="inferred">
27 <desc>What is your holiday budget?</desc>
28
29 <answer value="low">
30      <display>Low budget</display>
31      <belief>You have a low budget</belief>
32      <explain>You want a quality holiday but are sensitive to the cost. You will not
be shown holidays dearer than $5000.</explain>
33 </answer>
34 <answer value="high">
35      <display>High budget</display>
36      <belief>You have a high budget</belief>
37      <explain>You will be shown all holidays regardless of cost.</explain>
38 </answer>
39 </attr>
40
41 <attr id="HolidayQuiz" acquisition="quiz_score">
42      <desc>The total number of quiz questions you answered correctly.</desc>
43 </attr>
44
45 <attr id="probLikesBeach" acquisition="prob" >
46      <desc>The probability that you like the beach.</desc>
47 </attr>
48
49 </inferred>
50
```

```
51 <init id="HolidayQuiz"
52      cond="default"
53      value="0"
54      evidence="You have not yet attempted the quiz."/>
55
56 <init id="probLikesBeach"
57      cond="default"
58      value="50"
59      evidence="There is no evidence about this yet."/>
60
61 <init id="budget"
62      cond="(self.group=='single')"
63      value="high"
64      evidence="SASY assumed you have a high holiday budget because you are single."
/>
65
66 <init id="budget"
67      cond="(self.group=='couple')"
68      value="low"
69      evidence="SASY assumed you have a low holiday budget because you are a Couple
and are probably saving for a house or paying off a mortgage." />
70
71 </um>
```

This profile page defines four attributes that are to be stored in each user's user model: *group* (line 6), *budget* (line 26), *HolidayQuiz* (line 41) and *probLikesBeach* (line 45). However, only the *group* attribute is directly acquired and is part of the initial profile since it is enclosed by the initial tags on lines 4 and 20. The other attributes are all inferred attributes and will only be displayed to the user, the second and subsequent times the profile page is accessed, when the inferred values have been determined by the system according to rules defined by the *init* tags on lines 51, 56, 61 and 66.

The first time a user accesses their profile, the initial profile questionnaire is displayed as in Figure 4.3. The questionnaire presents the user with a single question *What best describes your situation?* To which the user selects an answer as *Single* or *Couple*. This question will set the value for the *group* attribute, which is used to determine the users marital status. In the ATML code, the question shown to the user in the questionnaire is defined by the *desc* tag on line 7. The *group* attribute is a fixed value data type which may be set to the *single* (line 8) or *couple* (line 13). Lines 10 and 15 state the beliefs that are stored by SASY and are displayed on a content page where the *group* attribute eligious ed personalisation. For example, the belief *You are single* is displayed on the right hand side of the page in Figure 3.4 because it caused content to be added or removed from the page by personalisation.

When the initial profile questionnaire is saved, SASY evaluates the profile initialisation rules on lines 51, 56, 61 and 66. The first two rules are defaults, the other two trigger only when a certain condition is met. When the profile is saved, the first rule (line 51) will set the *HolidayQuiz* to *0* and add the evidence description *You have not yet attempted the quiz*. This text is displayed by the SASY Evidence

tool that helps the user understand how their user model was populated. The last rule (line 66) sets the *budget* attribute to *low* if the *group* attribute is *couple*, as defined by the condition on line 67. That is, SASY will assume the couples have a low budget. This illustrates the use of Python code style expressions in ATML.

Rules, that is, *init* tags, are evaluated every time the profile is saved but will not override values specified by the user explicitly. For example, if the condition on line 67 evaluates as true, it will not update the budget attribute if the user has explicitly changed the budget attribute on the profile page because user input is considered stronger than inferred information.

The initial profile questionnaire does not show the inferred attributes, but does provide the automatically generated heading “So far, SASY has inferred the following about you”. This serves to inform the user that there will be some information that is inferred about them.

On subsequent visits to the profile page, it appears as in Figure 4.4, showing both initial and inferred user model attributes and their current values. The profile now shows the *group* (line 6), *budget* (line 26), *HolidayQuiz* (line 41) and *probLikesBeach* (line 45) attributes. In this example, the user has selected *Couple* in their initial profile. Notice that the rule on line 66 fired to cause SASY to build the assumption the user has a low budget.

The current value of each user model attribute is a hyperlink that invokes the Evidence tool when clicked. For example, clicking *Low budget* in the middle column opens the evidence describing how this attribute was set, as in Figure 3.8. Notice this evidence text was specified on line 69.

Another important aspect of the profile page when displayed by SASY is the third column of the page. The *explain* tags on lines 32 and 37 are rendered and displayed in this column, as shown in Figure 4.4. For example, line 32 causes the text *You want a quality holiday but are sensitive to the cost. You will not be shown holidays dearer than $5000* to be provided on the profile page to explain the effect of selecting a low budget. This text helps the user understand the impact of the different values for user model attributes.

Finally, Figure 4.4 shows that *quiz_score* attributes can be reset to zero by the user by checking the associated check box on the profile page. Similarly, *prob* attributes can be reset to 50 %.

4.5 Content Pages

We first present the tag structure of content pages, briefly describe the key tags and illustrate with an example. Details for each tag are provided in the appendix 8.1.2.

4.5.1 Document Structure

Each SASY topic contains one or more content pages. A content page may contain a mixture of static content (for example HTML tags) and adaptive content. Content pages may contain any combination of HTML tags and ATML tags. The following ATML tags are the discussed in this section:

- *adapt* tag – used to mark-up content as adaptive by specifying rules to conditionally display the enclosed content to the user if the values in their user model meet a certain criteria.
- *link* tag – used to create an annotated hyperlink for which the author can specify pre-requisite pages the user should access before following the link.
- *updateum* tag – used to update the user model attributes, either on access to a page or on a mouse click.
- *question* tag – used to generate quiz questions that update the user model when an answer is selected.

4.5.2 Adapt tag

The *adapt* tag marks content as adaptive. The tag specifies a condition that must be satisfied by a user's user model in order for that user to be shown the enclosed content in their personalised view of the page. It also specifies information that is shown to users when they scrutinise the personalisation to understand why adaptive content was added to or removed from a page.

The *adapt* tag may contain the same sub-elements as the page tag. For example, it may contain additional *adapt* tags and HTML tags.

This tag can be used in two forms: simple and complex. This is explained below.

Simple Adaptation

The simple form requires the *cond* and *value* attributes to be specified, as shown in the example below. The *cond* attribute references the identifier of a user model attribute (as defined by an *attr* tag on profile page) and *value* specifies the value that attribute must have for the enclosed content to be added to the user's personalised view of the page. For example, consider the following ATML code:

```
<adapt cond="budget" value="low">
Stay at Sydney Backpackers for only $30 a night
</adapt>
```

The content *Stay at Sydney Backpackers for only $30 a night* is only added to a user's personalised view of the page if their user model attribute *budget* has the value *low*. Otherwise the content is removed from their personalised view.

For this form of the *adapt* tag, scrutinisation information, displayed by the Highlight tool, is automatically generated by SASY that explains why the adaptive content was added to or omitted from the user's personalised view of the page. Following the above example, if the content is added the following text is provided by the Highlight tool "Added because profile has: You have a low budget" Likewise, it is was omitted the alternate text is shown "Removed because profile has: You have a low budget".

Complex Adaptation

The complex form requires the *cond*, *type*, *inc_explain* and *exc_explain* attributes to be specified. In this form, *type* is always set to the constant "complex". This form allows the author to defines a logic expression involving Python-style code (in the *cond* attribute), as described for the *init* tag earlier in this chapter.

Since the logic expression for adaptive content inclusion/omission is arbitrarily defined by the author, the author must also provide a textual explanation of the logic that will be shown to the user. The *inc_explain* and *exc_explain* attributes provide information that is displayed to the user when they scrutinise using the SASY Highlight tool:

- *inc_explain* – text shown to the user to explain why the enclosed content was added by personalisation.
- *exc_explain* – text shown to the user to explain why the enclosed content was removed by personalisation.

For example, consider the following ATML code:

```
<adapt type="complex" cond="string.atoi(self.HolidayQuiz) > 10"
inc_explain="You have answered enough questions correctly to pass the quiz."
Exc_explain="You have not answered enough questions correctly to pass the quiz."
>
You have passed the quiz
</adapt>
```

The content *You have passed the quiz* is only added to a user's personalised view of the page if their user model attribute *HolidayQuiz* indicates they have answered more than 10 questions correctly. Otherwise the content is omitted from their personalised view. The code `string.atoi()` is a standard Python function that converts a string to an integer.

If a user who correctly answered more than ten questions scrutinises the personalisation using the Highlight tool, the text *You have answered enough questions correctly to pass the quiz* is shown to them by the Highlight tool. For users who have not passed the quiz, the alternate text is shown *You have not answered enough questions correctly to pass the quiz*.

4.5.3 Link tag

The *link* tag provides an annotated hyperlink, inspired by the ELM-ART system. (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001). The hyperlink label is colour coded based on a traffic light metaphor. The author may specify zero or more pre-requisite pages using the *pre* attribute. If the user has visited all the pre-requisite pages, the hyperlink appears in green, indicating the user is ready to read the material. If the user has not visited all the pre-requisite pages, the hyperlink appears in red. If the user has already visited the page, its hyperlink will appear in black, but still underlined to distinguish it from normal text. Hyperlink colour coding indicates which pages the hypertext author considers the user should read or not, based on the current state of the user's user model. At the same time, the user is free to access any page.

An example of how this tag is rendered is provided in Figure 3.10 (pg. 90).

4.5.4 Updateum tag

This *updateum* tag is used to conditionally update the value of a user model attribute. It allows the user to specify rules to update the user model on content pages, similar to the *init* tag that is used to set initial user model attributes. The *updateum* tag is evaluated, and the user model is updated, when the page containing it is loaded. The position of the tag on the page is important: for example, if it is embedded at the top of an ATML page, it will update the user model before the remainder of the ATML page is processed. We describe the attributes below, an example is described in the next section.

The *attr* attribute identifies the user model attribute to be updated. The *value* attribute specifies the new value to be assigned. Alternatively, the *inc* attribute specifies the increment to be applied to the current value, that may be positive or negative. The *inc* attribute is used to update probability type user model attributes.

The *strength* attribute indicates the certainty of the update: *direct* means it is certain, *inferred* means it is not certain, perhaps based on an observation about the user. User model updates made by the user through the profile page are considered *direct*. So if the current value of the attribute is *direct*, only another *direct* update will be applied; *inferred* updates will be ignored because there is already strong (i.e. direct) evidence of the current value. This stops the adaptive inference functionality of the system from overwriting changes made to the user model by the user.

The *evidence* attribute is used for scrutability. This text is used to explain why the user model was updated. It is stored in the user model and displayed by accessing the Evidence Tool, that can be invoked either from the profile page, or a content page where personalisation occurred due to the user model attribute.

The *addEvidence* attribute specifies whether the evidence text is to be always added to the profile or only if the current value of the user model attribute is changed.

The optional *ifCond* attribute allows the author to specify a condition for whether the user model should be updated at all. The condition is defined through Python-style code, as for the *init* tag described earlier in this chapter. The user model is only updated if the expression evaluates to true. Otherwise the *updateum* tag is ignored. If no condition is specified, that is, no *ifCond* attribute is specified, the user model is always updated.

Update User Model on User Events

It is also possible to evaluate an *upateum* tag on a mouse click, or other user event, through custom JavaScript code. To do this, the author first crates a separate ATML document containing the *updateum* tag(s) that are to be evaluated on the event. For example, the author defines all rules in the new ATML document *eval.xml*. The author then adds JavaScript to the ATML document where the user event (e.g. mouse click) will occur. The JavaScript invokes SASY to process the new ATML document (eval.xml) in the background when the event occurs. This is done by specifying "log" as the target browser window to load *eval.xml* in to. This is a special hidden window in SASY, thus the page is processed in the background. For example, consider the following HTML and JavaScript code that could be added to a regular content ATML page:

```
<a href="get.cgi?page=Courses/HOLIDAY/eval.xml" target="log">Click Me</a>
```

This is rendered as a hyperlink by the browser. Clicking the link causes the page *Courses/HOLIDAY/eval.xml* to be loaded in to the target window, *log*, that is a special hidden window in SASY. All *updateum* tags on the *eval.xml* page will be processed in the background. There are also other approaches to perform this task using JavaScript, but only this simple form is presented here.

4.5.5 Content page Example

The following XML is an example of a typical content page. An example of how the page is rendered for a user who has a low budget and likes the beach is shown in Figure 4.5. Figure 4.6 shows the same page but with the Highlight tool invoked to allow the user to scrutinise the personalisation on the page.

The page tag on line 3, specifies that the *budget* and *probLikesBeach* user model attributes are used for personalisation on this page. This is rendered as a list on the right hand side of the

page, which shows the text *The probability that you like the beach is 60* and *You have a low budget*. Clicking the *Why?* link after these labels opens the Evidence tool to show evidence about how these user model beliefs were set. The Evidence tools was described in Chapter 3.

```
1 <?xml version="1.0"?>
2 <lesson>
3 <page umKeys="budget,probLikesBeach">
4
5 <updateum attr="probLikesBeach"
6      inc="10"
7      strength="inferred"
8      evidence="You accessed a page about beach holidays"
9      addEvidence="always"
10 />
11
12 <BR/><H3>Here is the perfect holiday escape for You!</H3>
13
14 <adapt cond="(string.atoi(self.probLikesBeach) > 50)" type="complex"
15 inc_explain="You like the beach"
16 exc_explain="You do not like the beach"
17 >
18
19 <P/><h4>Paradise in Aitutaki</h4>
20 Imagine sitting in a hammock, cocktail in hand, lolling under a palm tree
doesn't
21 sound too bad at all, eh? Life on Aitutaki, the Cook Islands second most
populour
22 island, is even more relaxed than on Rarotonga. It has a road that runs
around
23 the island close to the coast, passing through several villages. The round
trip
24 takes only a few hours.
25
26 <P/>
27 <img src="Courses/HOLIDAY/ait1.jpg" width="100" />
28 <img src="Courses/HOLIDAY/ait2.jpg" width="100" />
29 <img src="Courses/HOLIDAY/ait3.jpg" width="100" />
30
31 <adapt cond="budget" value="high">
32 <P/>The Aitutaki Lagoon Resort is the premier property on the island. The
Deluxe
33 bungalows have raised floors to maximise the view of the lagoon. There are
two
34 restaurants, a pool with a bar, boutique and activities such as snorkelling.
35 Prices start at $5000 per couple per week, but it is well worth the luxury.
36 </adapt>
37
```

```
38 <adapt cond=”budget” value=”low”>
39 <P/>Parasise Cove is well priced and is on the north-west side of the island
on
40 a white sandy beach. Each bungalow has a queen or two single beds and
refrigerator.
41 All other facilities are shared. Prices start at about $200 per person per
week.
42 </adapt>
43
44 </adapt>
45
46 <pagelink>
47     <prev auto=”true”/>
48     <next src =”Courses/HOLIDAY/page2.xml” title=”More holidays” />
49 </pagelink>
50
51 </page>
52 </lesson>
```

On line 5, the *updateum* tag is used to update the user model attribute probLikesBeach. The tag increments the current value by 10 and adds the following evidence text to explain this update *You accessed a page about beach holidays*. This text will be shown by the Evidence tool when the user accesses it.

Line 12 includes HTML code that defines a heading.

There are three adapted parts on this page, each enclosed by a pair of *adapt* tags. The first encapsulates content between lines 14 and 44, the second contains encapsulates lines 31 through 36, the third encapsulates content lines 38 through 42.

The first *adapt* tag, on line 14 uses the complex form of the tag. The effect of this tag is that the page content between lines 19 and 44, where the close tag is located, will only be shown to a user if the probability that they like the beach (i.e. the value of their user model attribute *probLikesBeach*) is more than 50. If it is included, the Highlight tool will show the text *You like the beach*. For example, see the upper part of Figure 4.6. Notice the text is displayed in blue font at the top of the grey area, just underneath the heading of the page.

Lines 19 to 29 include more HTML code.

Line 31 shows the use of the simple form of the *adapt* tag. The effect of this tag is that the enclosed content from lines 32 to 36 is only shown to users that have a high budget. That is, their user model has the value *high* for the user model attribute *budget*. On scrutinising through the highlight tool, the scrutinisation text is automatically generated by SASY for all simple adapt tags, so the author of the SASY document has less work to do for this form of. For example, in Figure 4.6, underneath the three

images in the yellow area, the page has the text *Removed because profile has: You have a low budget*. In this case, the user had a low budget and the paragraph describing the Aitutaki Lagoon Resort accommodation was removed from their personalised view of the page. Notice also that because this second *adapt* tag is nested within an *adapt* tag, the yellow area has a black border that is indented from the left and right edge of the grey area. This indentation helps describe the nesting of adaptive content sections.

Lines 46 to 49 include a *pagelink* tag. The *prev* tag on line 47 is rendered as left arrow at the bottom of Figure 4.5 while the *next* tag on line 48 is rendered as right arrow. The *prev* tag uses the auto attribute that instructs SASY to automatically generate a link to the previous page the user accessed, according to the page history stored in their user model.

Figure 4.5 – Content page that has been personalised to a user that has a low budget and likes the beach.

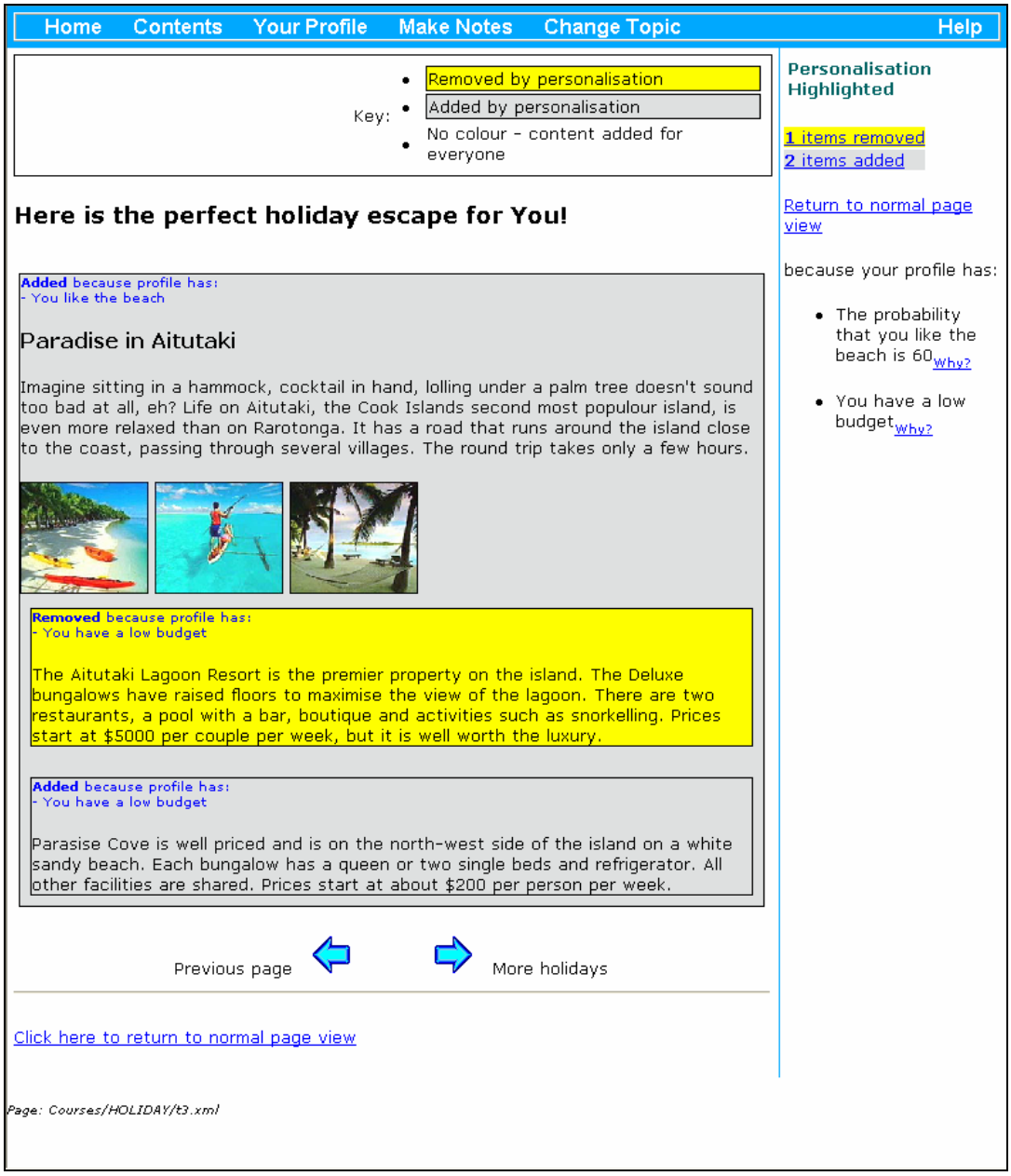


Figure 4.6 – Shows the same page as in Figure 4.5 but with Highlight Tool invoked to allow the user to scrutinise the personalisation.

4.5.6 Question tag

The *question* tag is used to create an interactive multiple choice quiz question using Dynamic HTML and JavaScript code generated by SASY. This tag is used on regular content pages, but is quite complex so is discussed in this section separately.

The author defines the question, answers, feedback given when and answer is selected. If the user selects the correct answer, a tick icon is shown and the corresponding feedback, as in Figure 4.9. If an incorrect answer is selected a cross icon is shown along with the corresponding feedback. The author can specify the maximum number of attempts the user is allowed. The author can specify rules to

update the user model when an answer is selected. The user model may be updated in several ways when an answer is selected:

- A score user model attribute (*quiz_score*) may be incremented when the question is answered correctly (see *scoreattr* attribute of *choices* tag).
- A probability (*prob*) user model attribute may be updated to increase or reduce the confidence level that the user knows a concept (see *concept* attribute of *choices* tag).
- An external ATML page may be processed in the background. This page may contain rules to update the user model using the *umupdate* tag (see *umupdate* attribute of *choice* tag). Unlike the first bullet point above, this option can be used to update multiple user model attributes or perform more complicated update logic. For example, *umupdate* tag could be used to check probabilities for certain knowledge concepts and update the user model so that the system presents remedial teaching material for areas in which the student has lacks competence.

Figure 4.7 shows the structure of the question elements. Full details provided in the appendix 8.1.2, here we describe the main features. The *choices* tag specifies which of the answers is correct, names the user model attribute used to keep the score (*scoreattr*), specifies the number of attempts allowed (*attempts*), and optionally the name of the probability attribute that is updated when an answer is selected (*concept*).

The *question* tag requires one or more *choice* tags. Each *choice* tag defines a possible answer to the question the user may select on the SASY page. The optional *umupdate* attribute of the choice tag specifies a URL for an ATML page that is processed when the answer is selected. This additional ATML page typically contains *umupdate* tags to update the user model but no content as this page is processed in the background and not displayed to the user. The optional *prob* attribute specifies an integer value (positive or negative), that is to be added to the user model attribute (defined in the *concept* attribute of the *choices* tag) when the answer is selected by the user. The *choice* tag requires *value* and *feedback* tags to be specified. The *value* is the answer that is selected by the user while *feedback* is the feedback message that is popped-up by the system when the user selects the answer.

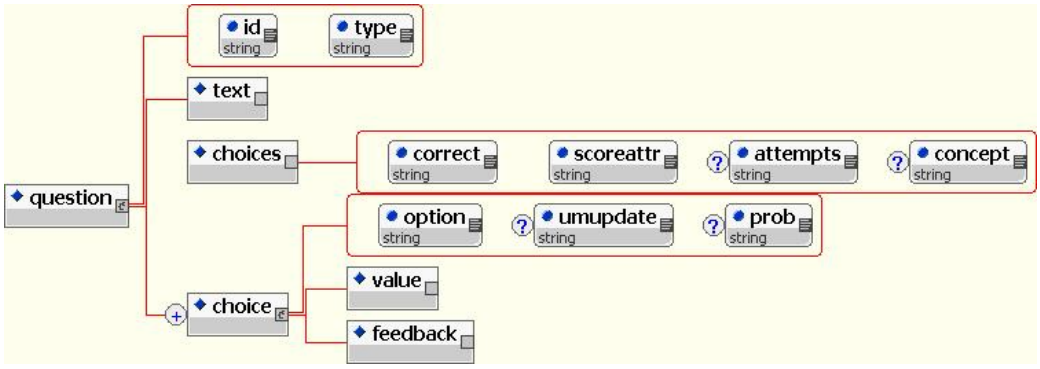


Figure 4.7 – Question tag structure

4.5.7 Quiz Question Examples

The following XML demonstrates the use of the question tag to generate interactive, multiple choice quiz questions. The rendered page is shown in Figure 4.8.

```
1 <?xml version="1.0"?>
2 <lesson>
3 <page umKeys="">
4
5 Answer this question correctly for your chance to win a free holiday!
6 <P/>
7
8 <question type="choice" id="Q1">
9 <text>
10 <P/>
11 What is the largest island of the Cook Islands?
12 </text>
13
14 <choices correct="1" scoreattr="HolidayQuiz" attempts="1">
15
16 <choice option="1" umupdate="Courses/HOLIDAY/eval.xml">
17 <value>Rarotonga</value>
18 <feedback>Correct.</feedback>
19 </choice>
20
21 <br/>
22 <choice option="2">
23 <value>Aitutaki</value>
24 <feedback>Incorrect. The largest island is Rarotonga.</feedback>
25 </choice>
26
27 </choices>
28 </question>
29
30 </page>
31 </lesson>
```

The *question* tag starts on line 8. The actual question is provided in the *text* tag on lines 9 to 12.

Line 14 specifies that the choice 1 is the correct one, the *HolidayQuiz* user model attribute is to be updated when the user answers the question and that users are only allowed one attempt at answering the question. Should they attempt to answer the question more than once, the following message is displayed *You have exceeded the maximum allowed number of attempts*.

Lines 16 through 19 provide the text and feedback for the first option, *Rarotonga*. The *value* tag is rendered as an HTML radio button. Clicking the radio button pops up the text supplied by the *feedback* tag on line 18. An example of how the feedback is rendered is shown in Figure 4.9.

As specified on line 16, when option 1 is selected, the ATML page *Courses/HOLIDAY/eval.xml* is processed in the background. This ATML page will not be displayed to the user, but might contain *updateum* tags to update some other attributes in the user model.

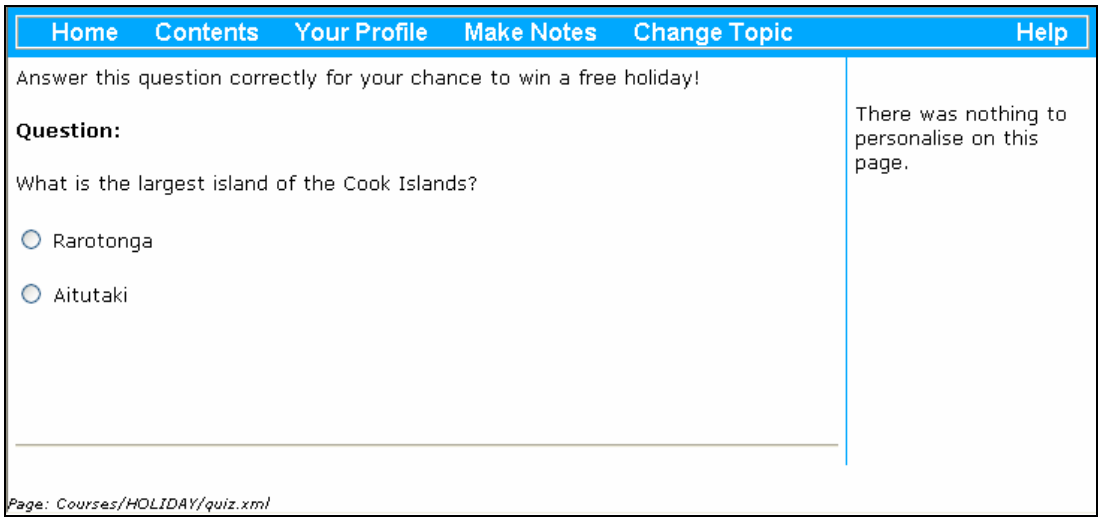


Figure 4.8 – Shows how a question tag is rendered in SASY as an interactive, multiple choice quiz.

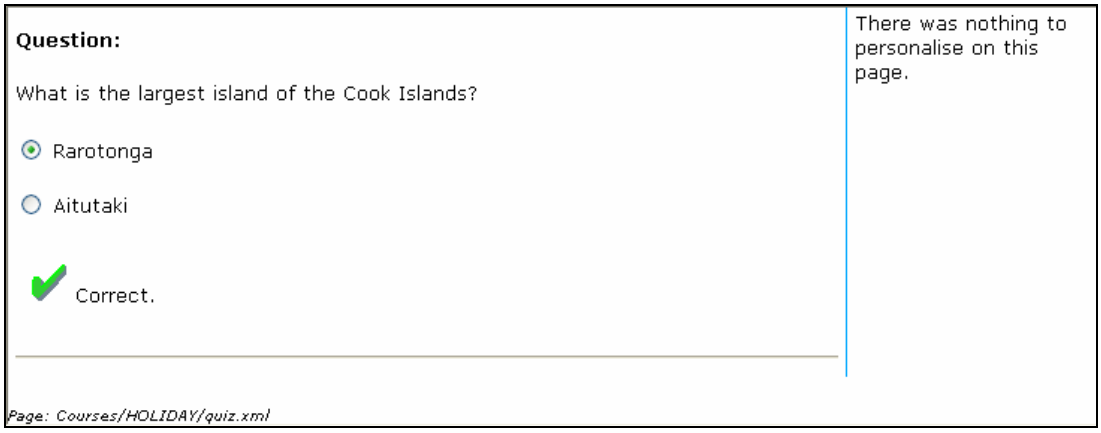


Figure 4.9 – Shows the feedback provided by SASY when the question in Figure 4.8 is answered by the user.

The following XML demonstrates how probability user model attributes can be updated through questions. Consider the following XML. Line 7 references the *probLikesBeach* user model attribute through the *concept* tag attribute. Lines 9 and 15 specify the probability that is to be added to the *probLikesBeach* attribute when the corresponding answer is selected. For example, if *O'ahu* is selected, the probability the user likes the beach is increased by 5.

```
1 <question type="choice" id="Q2">
2 <text>
3 <P/>
4 On which Hawaian island is Waikiki beach found?
5 </text>
6
7 <choices correct="1" scoreattr="HolidayQuiz" concept="probLikesBeach">
8
9 <choice option="1" prob="5">
10 <value>O'ahu</value>
11 <feedback>Correct.</feedback>
12 </choice>
13
14 <br/>
15 <choice option="2" prob="-5">
16 <value>Maui</value>
17 <feedback>Incorrect.</feedback>
18 </choice>
19
20 </choices>
21 </question>
```

4.6 Home Page

The home page has the special property that it is the first page displayed when the user enters a topic following the profile page.

Clicking *Home* in the menu at the top of any SASY page loads the Home page.

The home page is defined using the same mark-up as a regular content page, but the ATML document must be named *teacher.xml*.

4.7 Contents Page

Clicking *Contents* in the menu at the top of any SASY page loads the Contents page.

The home page is defined using the same mark-up as a regular content page, but the ATML document must be named *coursemap.xml*.

It is intended that the contents page be used to list all the pages that can be accessed in the topic, using the *link* tag to provide colour coded annotation, a form of adaptive guidance.

4.8 Summary

This chapter described the Adaptive Tutorial Mark-up Language (ATML), SASY 4. SASY content is created by defining a set of ATML documents, collectively called a topic.

This chapter first introduced ATML and the standard documents that must be created for each topic. The structure of ATML documents was discussed in detail.

The next chapter describes the evolution of the scrutinisation tools, to their form in SASY 4.

Chapter 5 Exploration of Tools for Scrutability Support

A central contribution of this thesis is the set of SASY scrutability support tools. They were created and evaluated in several stages. The results of each evaluation were analysed to reveal enhancements that could be applied to the tools. This chapter describes the evolution of the Scrutability Support tools that were developed.

The goal was to design and develop tools for scrutability support that would satisfy the requirements and described in comprehensively in Chapter 3, section 3.2 – page 75. To so this required addressing the challenges presented in section 3.3 – page 77.

This chapter describes the evolution of scrutability support tools developed and evaluated as part of this thesis: Tutor (versions 1, 2 and 3), Cell-Tutor and SASY 1. For each system we present the design goals, a brief system overview, an informal report of our evaluation of the system and conclusions drawn from this.

5.1 Tutor 1

5.1.1 Design Goals

Tutor 1, created in 1998, was the first version of the Scrutability support tools for adaptive hypertext. In this thesis, we deal with just the scrutinisation of Tutor 1, which is fully described elsewhere (Czarkowski 1998; Czarkowski & Kay 2000).

The design goals for this version were to satisfy the requirements described in Chapters 1 and 3, starting with a simple form of scrutinisation tools. Personalisation was driven by logic conditions embedded in each web page and a simple user model, defined entirely by the user through a brief online questionnaire. Although the approach was simple, it made a logical starting point for exploring ways to effectively provide support for users to scrutinise the adaptation of a system.

Tutor 1 was also developed with a focus on delivering personalised teaching content. Most of these features have been carried forward to SASY.

5.1.2 System Overview

The system architecture is identical to SASY 4, as described in Chapter 3. The key differences from SASY 4 lie in the scrutability tools, complexity of personalisation and how the user model is updated.

The interface design and system features were strongly influenced by the ELM-ART system which teaches LISP (Weber & Brusilovsky, 2001).

To use Tutor 1, users must first register for a user account and then log in. Having logged in, users can select from several online courses. On starting a learning topic, users are presented with a questionnaire that is used to establish their profile (user model). The questions asked, determined by the hypertext author used, are used to populate the user model. For example, *Would you like to be asked questions to test your understanding?* With corresponding possible answers like *Yes* or *No*. Unlike SASY, the user model is static and is not updated or changed by the system.

Having created an initial profile, users see the main Tutor 1 interface, such as in Figure 5.1. The icons across the top of the page are hyperlinks to various teaching facilities:

- Course map – same as SASY course map page.
- Teacher's news page – same as SASY course home page.
- Notes editor – same as SASY notes editor.
- Glossary – an adaptive page that provided definitions of terms used in the course.
- Profile – similar to SASY Profile page.
- Discussion forum – a bulletin board, chat room, facility where students could post questions and messages to the community of Tutor 1 users.
- Log out – logged out of Tutor 1.
- Help – displayed online help.

Tutor 1 had two scrutinisation tools: Profile and Adaptation Explanation.

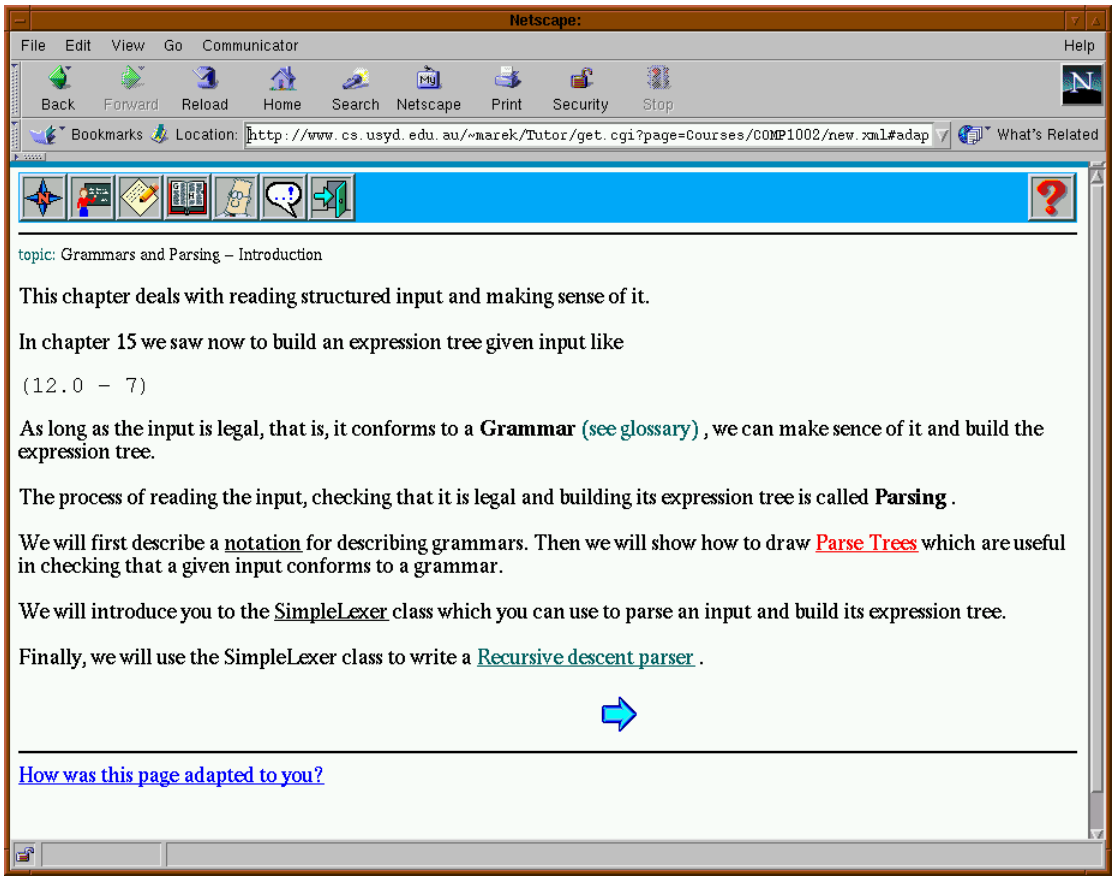


Figure 5.1 – Example of adapted teaching page in Tutor 1.

The Profile allows the user to view and change their current user model attributes. The Profile is similar to SASY, except in Tutor 1 the profile does not provide an explanation as to why a profile attribute was set (in SASY the Explanation tool can be accessed from the profile page by clicking on any profile attribute). However, there was no need to provide explanations since the system did not update the profile. The user was under the direct and exclusive control of the user. At any stage, the user can see their profile and alter it.

The purpose of the Adaptation Explanation tool is to inform the user why content on a page had been adapted. This tool is the predecessor of SASY's Highlight tool. At the bottom of every teaching page is a link, *How was this page adapted to you?* As in Figure 5.1, this is below the main navigation arrows. The link appears in blue, underlined font, distinguishing it from other hyperlinks that were colour coded in black, green and red font: denoting links that have been visited by the user, are ready to be visited, are not ready to be visited because the user does not have enough pre-requisite knowledge.

When the user clicks the *How was this page adapted to you?* Link, the bottom of the page expands to include a summary of the personalisation on that page, as in Figure 5.2. If there was no personalisation, the following message is displayed instead: *There was no adaptation of this page*.

The Adaptation Explanation includes a separate explanation for each personalisation (i.e. adaptation) that occurred, including:

- the message *excluded* (if the content was removed by personalisation) or *included* (if the content was added by personalisation);
- the profile question that caused the content to be included or excluded (*Condition*);
- the answer to the profile question that would cause the content to be included (*Inclusion value*);
- the user's current answer to the profile question (*Your value*).

For example, the explanation in Figure 5.2 shows there were two adaptations performed on the page. The first caused content to be removed from the user's view of the page because the user answered *no* to the profile question *Would you like to see suggestions for exercise questions?* The second caused content to be added because the user answered *yes* to the profile question *Would you like to be asked questions to test your understanding?* In both, the *Inclusion value* indicates the answer that was required for the content to be included, while *Your value* shows the user's current answer.

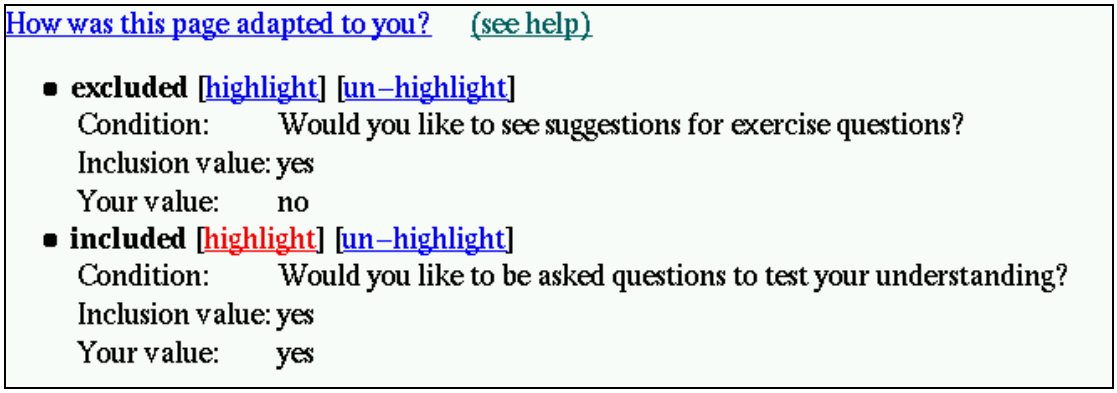


Figure 5.2 – Example of the Adaptation Explanation tool in Tutor 1.

If the user clicks on he text *[highlight]* for an aspect that was included, as in the second case in Figure 5.2, Tutor 1 highlights the background of the relevant material in yellow, similar to SASY, and moves the screen display to the point where the material is positioned on the page. An example of the result is shown in Figure 5.3.

If the user clicks on the text *[highlight]* where adaptation involved omission of material, as in the first case in Figure 5.2, Tutor 1 moves the screen display to the point where the additional material might have been placed and shows the message *...content removed here....* An example is shown in Figure 5.4. Although it would have been quite feasible to show the actual content that had been omitted, it was thought that it is easy for the user to alter their profile and see how that affects the document. Once the user has examined the explanations of the adaptations, they can click the *[hide explanation]* link at the bottom of the screen, causing the display to revert to the usual form as in Figure 5.1. This reduces the clutter of the explanations and the distraction from the main learning task.

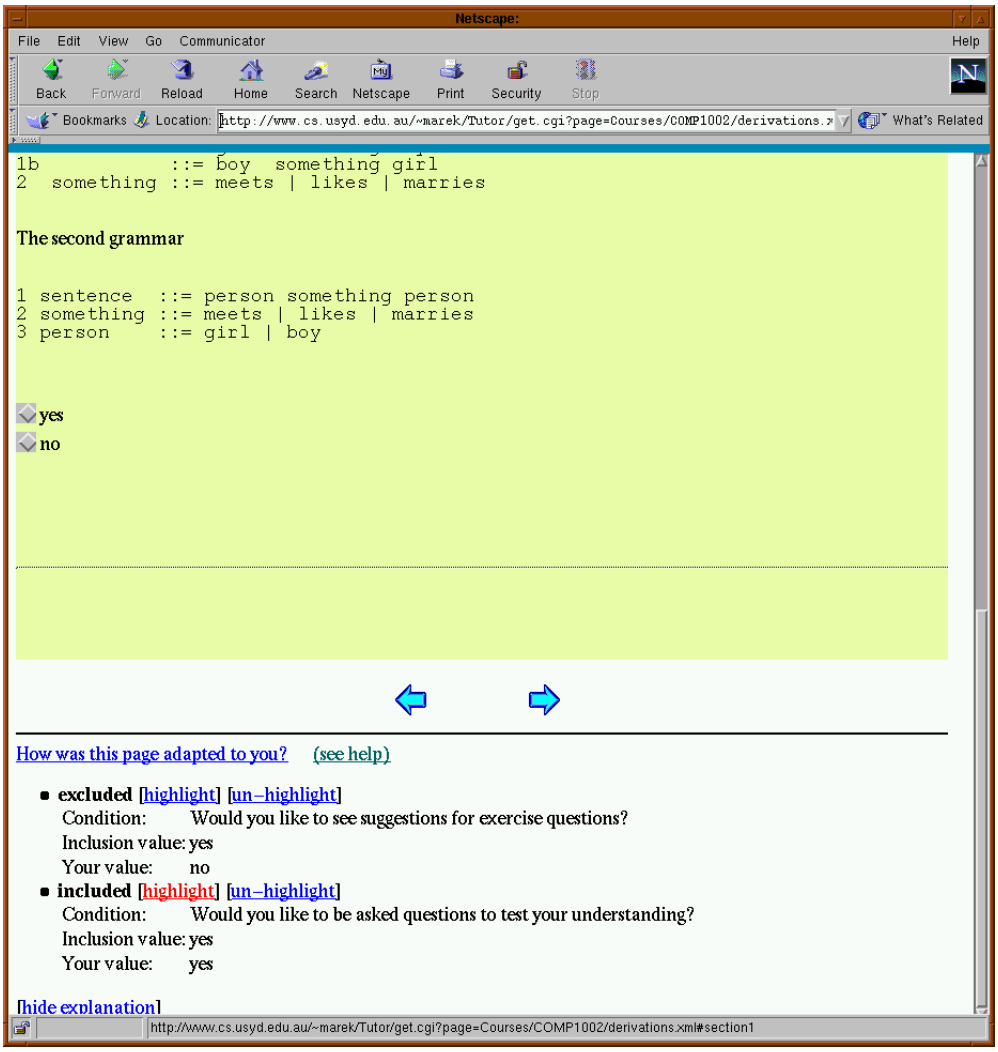


Figure 5.3 – Example of the content included by personalisation being highlighted from the adaptation explanation.

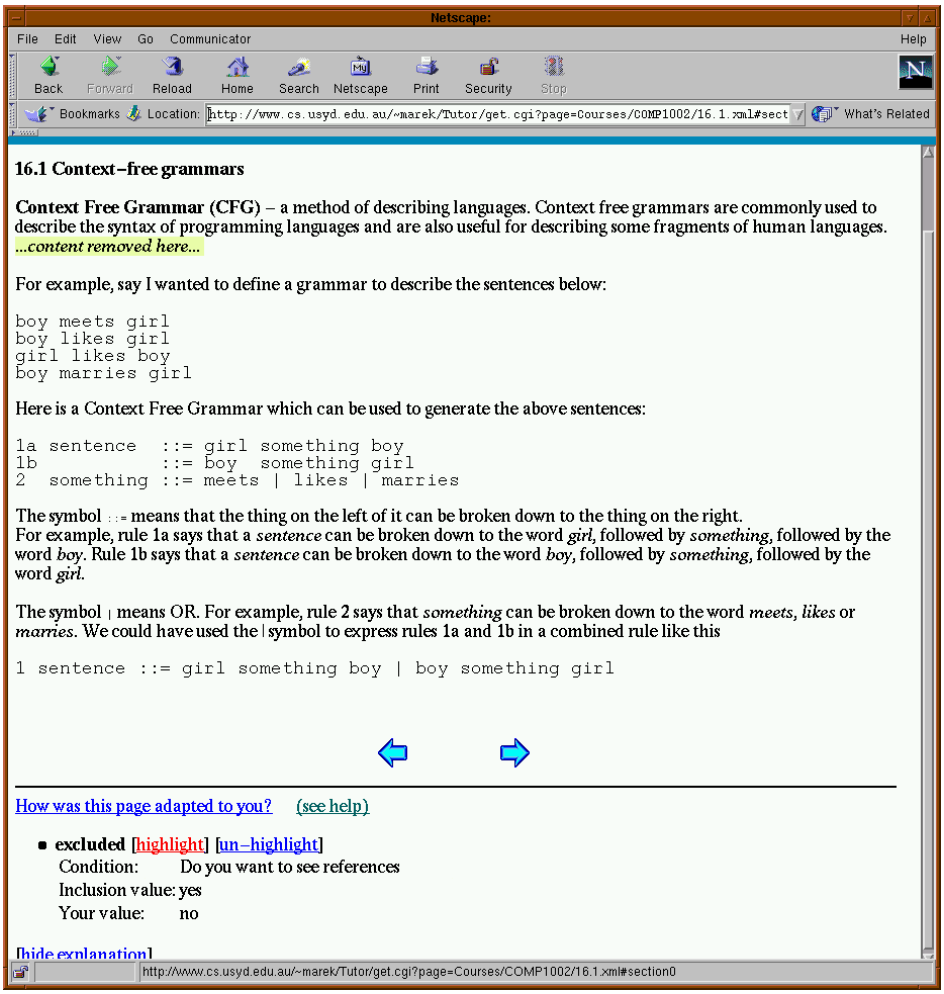


Figure 5.4 – Example of the content removed by personalisation being highlighted from the adaptation explanation. First paragraph includes text *content removed here*.

5.1.3 Evaluation

As described in (Czarkowski & Kay 2000), we conducted a field trial of Tutor 1 in the second semester of a large first year course in 1998 at the University of Sydney, Australia. The whole class of about 600 students was mailed an invitation to try Tutor 1 to learn about UNIX. The system logged user actions. The system encouraged users to use their system login name for Tutor 1 but there was no checking possible. In all 114 users registered, ostensibly 108 as students in the course and 6 as teaching staff. A summary of log data is shown in Table 5.1

Table 5.1 Summary of Log data for Tutor 1 field trial.

N=108	**Average**	**Median**	**STD range**	**Range**	**Percentage of larger use**
Logins	3.5	2	3.4	0-18	65 > once
Use of Profile	1.4	1	1.1	0-7	28 > once
Use of Adaptation Explanations	1.0	1	3.0	0-22	29 >= once

The first row indicates that the average number of logins per user was 3.5, with the median being 2 and the standard deviation 3.4. The range of logins was 0 to 18 and 65 % of users used it more than once. Inspection of the log indicates some of this activity was due to problems with the browser where cookies were not enabled. Also, some students were regularly checking the *Discussion Room*. The zero values correspond to users who registered but never logged in (mainly because of browser problems or disabled cookies).

The second row shows the number of accesses to the profile page. All students who made use of the teaching material had at least one access to it as the system automatically brought up the profile page. However, the log indicates that 28 % of users returned to the profile at least once more.

Analysis of the student profiles indicates that two types of student stereotype emerged. The majority matched a stereotype which preferred all the optional material (85% wanted to be asked quiz questions, 85% wanted easy/lead up material, 80% wanted references to background information), while a minority stereotype elected for the basic minimum form of the material.

The final row of the table shows the level of use of the explanations of the adaptation. This facility was clearly not used heavily. However, 29% of users made at least once use of it. Examining the logs, it turns out that many of these selections were for pages which had no adaptation. It is possible that if users took the effort to click the link only to be shown the message *There was no adaptation of this page*, this may have discouraged users from exploring the explanation again. Of those who sought explanations on a page that had been personalised, 27 % went on to select a *[highlight]* link to see what had been adapted.

The Tutor 1 interface also allowed users to leave free-form comments. Overall, these were few but very positive about Tutor 1, the teaching materials, the adaptivity and the support for scrutability. One user commented that the traffic-light colour coding of hyperlinks helped him to navigate his way through

the course material. He explained that he did not have a good grasp of the English language and so it would have been difficult to decide which links to follow based solely on their textual labels.

Another user commented that it was difficult to see exactly what had been removed by the personalisation. To her, seeing what had been removed was more important than seeing what had been included because she wanted to make sure she had not missed out on learning something important.

5.1.4 Conclusion

In our initial trials of the system, a substantial minority of students explored the explanations of adaptation: 29% examined the adaptivity and 27% of these users went further, checking the actual parts of the hypertext page affected by the adaptation. This level of interest is quite high, given that the students saw our system as offering additional assistance for a challenging part of the current course work. Their focus was, naturally, on learning. Any time devoted to exploring the adaptation was a distraction from that main task. The study showed that some of these users were definitely interested in scrutinising the adaptive hypertext.

User feedback indicated the system had a limitation. Users could not easily see what content had been removed by personalisation other than by changing their profile to have the removed content included instead.

However, this field study did not provide enough information to conclude whether the scrutinisation tools were effective. The next step was to conduct a qualitative study.

5.2 Tutor 2

The logging study of Tutor 1 showed some users were keen to scrutinise. The next step was to carry out a qualitative to examine user comprehension of the adaptation that occurred and how they, as the reader in control, could control it to suit their needs. At the same time, we identified some interface enhancements in the evaluation of Tutor 1. These are discussed in the following section.

5.2.1 Design Goals

There was one new design goal: to enable users to easily see what content had been removed by personalisation. This was driven by feedback received about Tutor 1 that users wanted an easy way to see all the content that a page offered, essentially a non-personalised version. This would allow them to see all the available content including that which had been removed by personalisation. The Tutor 2 version of the scrutability support tools are also described in Czarkowski & Kay (2001).

5.2.2 System Overview

We made several modifications to the Adaptation Explanation that appeared at the bottom of a page when the user clicked the link *How was this page adapted to you?* The revised form is shown in Figure 5.5. There are several differences from the Tutor 1 version in Figure 5.1.

Firstly, the Tutor 2 explanation stated whether it was content or navigation that had been adapted. In Figure 5.5 the first explanation states content was included, the second states a navigation element was excluded. It was intended this annotation would be useful in describing the content that had been adapted.

Secondly, since Tutor 1 had positive feedback about the colour coding of hyperlinks, the adaptation explanations links in Tutor 2 were also colour coded, following the same traffic life analogy. Explanations that referred to included content items were coloured in green, explanations for excluded items were in red font. For example, in Figure 5.5, the first bullet point text "page content included" is in green font, the second bullet point text "page navigation excluded" is red.

Thirdly, the explanation section included another hyperlink, *Show all text*. Clicking this reloaded the page and displayed all content available for the page, ignoring the adaptation. Items that would be included on adaptation were displayed in a green background (i.e. highlighted), items that would be excluded had a red background. This feature was intended to allow the user to more easily see the content that had been removed by personalisation.

Finally, *Your Value* showed the user their current answer to a particular profile question. In Tutor 2, this was a hyperlink to the profile page, making it easier to navigate to the profile page to change an answer to a profile question.

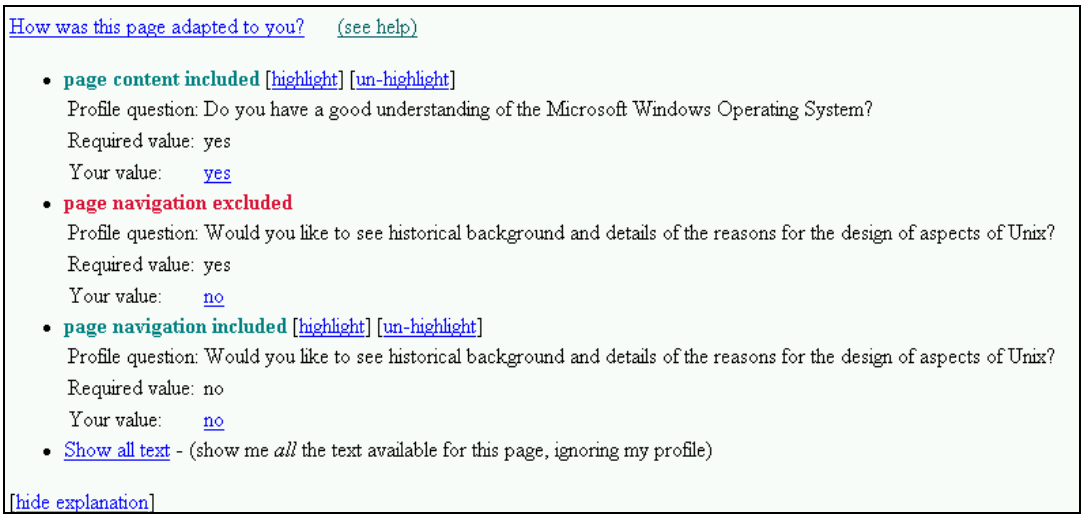


Figure 5.5 – Adaptation Explanation in Tutor 2.

5.2.3 Evaluation

An informal, qualitative evaluation was conducted in January 2002. The aim of the evaluation was to obtain feedback about the usability and user's first impressions of the scrutinisation tools; to observe whether users could in fact use them to inspect and control the personalisation.

An interviewer was hired to observe and interview participants as they used the system. Participants were asked to think-aloud as they used the system. Their actions were observed and comments were recorded (included in Appendix 8.2, page 229). We selected four first year computer science students from the University of Sydney since UNIX was part of the first year syllabus and we hoped they would be motivated to use Tutor 2 as a serious learning tool. None of the participants had prior experience with UNIX.

Two of the participants were given a short briefing on how to use the system. The introduction included a walkthrough of the adaptation explanation functionality of Tutor 2. The other two participants were not trained. All users had access to online documentation that briefly described how Tutor 2 adapted course material and how to change the user model to alter adaptation. However, none of the participants accessed the help. We wanted to observe differences in the responses between the trained, not-trained participant groups would indicate whether training helped users to understand and use the adaptation explanations.

Of the two untrained participants, one was able to fully grasp how Tutor 2 was adapting the course material and how they could control it. The other participant had the misconception that the purpose of the adaptation explanation was to provide hyperlinks to additional related course material. This participant also thought the highlight feature was a mechanism the annotate important information on the page. Having been prompted by the interviewer to try out and explain the links in the adaptation explanation, the participant eventually worked out the purpose of the adaptation explanation and that adaptive content was included if the criteria for the content matched their profile.

Participants who were trained did understand the adaptation explanations, but only one participant appreciated the level of control which they had over the adaptation. For example, one participant commented the amount of course material provided by the system was overwhelming. But when asked whether they considered changing their profile to change the personalisation to reduce the amount of content displayed, the participant said they had not considered this but would now that it had been suggested. The participant initially thought that only examples and quiz questions were adapted but not the other content. The other participant clearly understood the personalisation and control they had over it.

All users expressed difficulty in reading and using the adaptation explanations, stating that the explanations were too long and overwhelming. The purpose of the adaptation explanations was not intuitive and participants required training or significant practice to develop competence using the tool.

This indicated adaptation explanations might be effective for expert users but were not adequate for novice users.

Not all users appreciated that they had complete control over the personalisation nor understood how to exercise this control.

One user commented that the link *How was this page adapted to you?* Was both inappropriately titled and hidden. The user stated they would not think to click this link because it looked like a link to information about the system, rather than something useful. However, we did not regard this comment as indicative of the general user perception.

The evaluation was informal and the interviewer indeed had a lot of dialog with the participants, asking them to explain what they thought they saw and prompting them to try the system features. However, we wanted to know how users perceived the scrutinisation tools. Thus, they were prompted to use the tools in order to prompt comments and feedback from participants. The limitation of this evaluation approach was that we could not ascertain whether users would have found and accessed the scrutinisation tools on their own.

5.2.4 Conclusion

The qualitative evaluation made it apparent that users found it difficult to inspect and change how pages had been personalised. It was apparent the participants did not initially understand the purpose of the adaptation explanation.

We discussed ideas for improvements to the interface with users. A common point emerged from participant feedback that there was too much information presented in the adaptation explanation and that perhaps it should be summarised into a condensed form.

5.3 Tutor 3

The qualitative evaluation of Tutor 2 showed users could not effectively scrutinise. The adaptation explanation in Tutor 3 was completely redesigned, based on the design goals presented in the next section.

The evaluation method was also redesigned to evaluate the scrutinisation tools in a more genuine environment, where the user was self-motivated to scrutinise rather than being prompted to access the tools. Also, we designed the evaluation to allow participant results to be directly compared. In the evaluation of Tutor 2, participants essentially had free reign of the system. This meant they saw different things at different stages in the evaluation, making it difficult to compare their experiences.

5.3.1 Design Goals

The evaluation of Tutor 2 indicated that users could not do the basic scrutability tasks we had defined as essential. So, in light of these findings, we completely redesigned the adaptation explanation.

The new design goals were:

- Present the adaptation explanation more concisely, in a summarised form that the user can further scrutinise if they wish.
- Provide a consistent method for viewing included and excluded adapted content.
- Add instructional text to the interface for the scrutinisation tools to inform the user how to do things and explain how the adaptation works.

5.3.2 System Overview

This section describes only the changes Tutor 3 from Tutor 2. Firstly, we added instructional text throughout the system to introduce and explain the scrutinisation tools. This was done because the evaluation of Tutor 2 revealed these the tools and the concept of scrutinisation is new and not immediately understood. On logging in for the first time, the system presented a welcome message explaining the scrutinisation tools:

Tutor is designed to adapt its teaching material to you. Each lesson will be adapted specifically to you, in order to match your unique background.

To do this, Tutor needs to learn a little about you first.

On the next page, you will be asked to answer a few questions about yourself. This information will be kept confidential. Only yourself, your teacher and Tutor will have access to this information.

The answers you give will be stored by the system, in what Tutor calls your profile, and will help Tutor adapt each lesson to you.

You will be able to review and change your profile at any time and, by doing so, alter how the material is adapted to you. At the bottom of every lesson page, there is an explanation of how the lesson has been adapted to you.

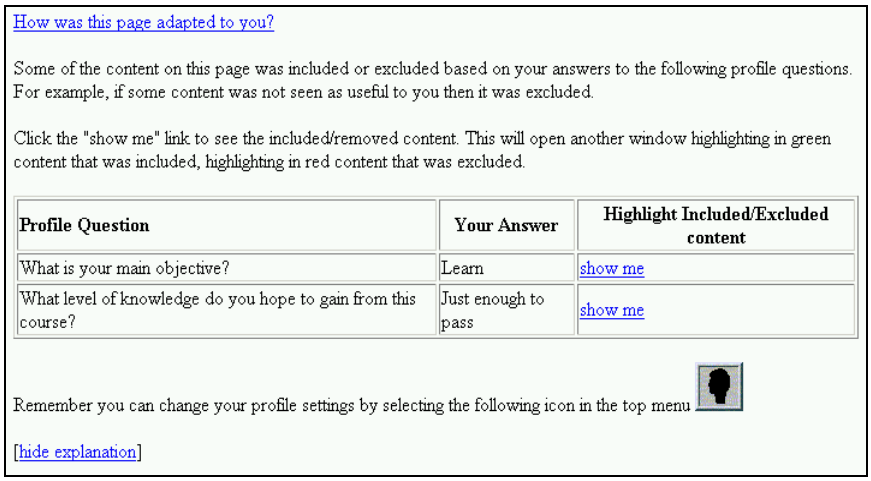
We also added the following text to the top of the profile page that was displayed every time the user accessed their profile:

Fill in your profile to suit your background, learning preferences, interests and current goals. Tutor uses your answers to adapt the content to your need. You can change your profile at any time during the course to influence Tutor's adaptation.

Finally, we added the four lines of instructional text to the top of the adaptation summary, shown in Figure 5.6.

On pages where content has been adaptively included or excluded, Tutor dynamically included the hyperlink *How was this page adapted to you?* At the bottom of the page. On pages with no adaptation the text *There was no adaptation on this page* appears instead of the link. This meant users would not need to click the link only to reveal that in fact there was no adaptation.

Clicking *How was this page adapted to you?* Displays an adaptation summary, just below the link, as in Figure 5.6. This summary is the starting point for scrutinising the adaptation of the current page. It tells the user which elements from their profile caused the adaptation. We called this the Adaptation Summary rather than the Adaptation Explanation, as in previous Tutor versions, because it summarises what personalisation occurred on the page but does not explain why it occurred. From the summary, the user can probe even deeper to see specifically what has been included or excluded because of their profile.


How was this page adapted to you?

Some of the content on this page was included or excluded based on your answers to the following profile questions. For example, if some content was not seen as useful to you then it was excluded.

Click the "show me" link to see the included/removed content. This will open another window highlighting in green content that was included, highlighting in red content that was excluded.

Profile Question	Your Answer	Highlight Included/Excluded content
What is your main objective?	Learn	show me
What level of knowledge do you hope to gain from this course?	Just enough to pass	show me

Remember you can change your profile settings by selecting the following icon in the top menu

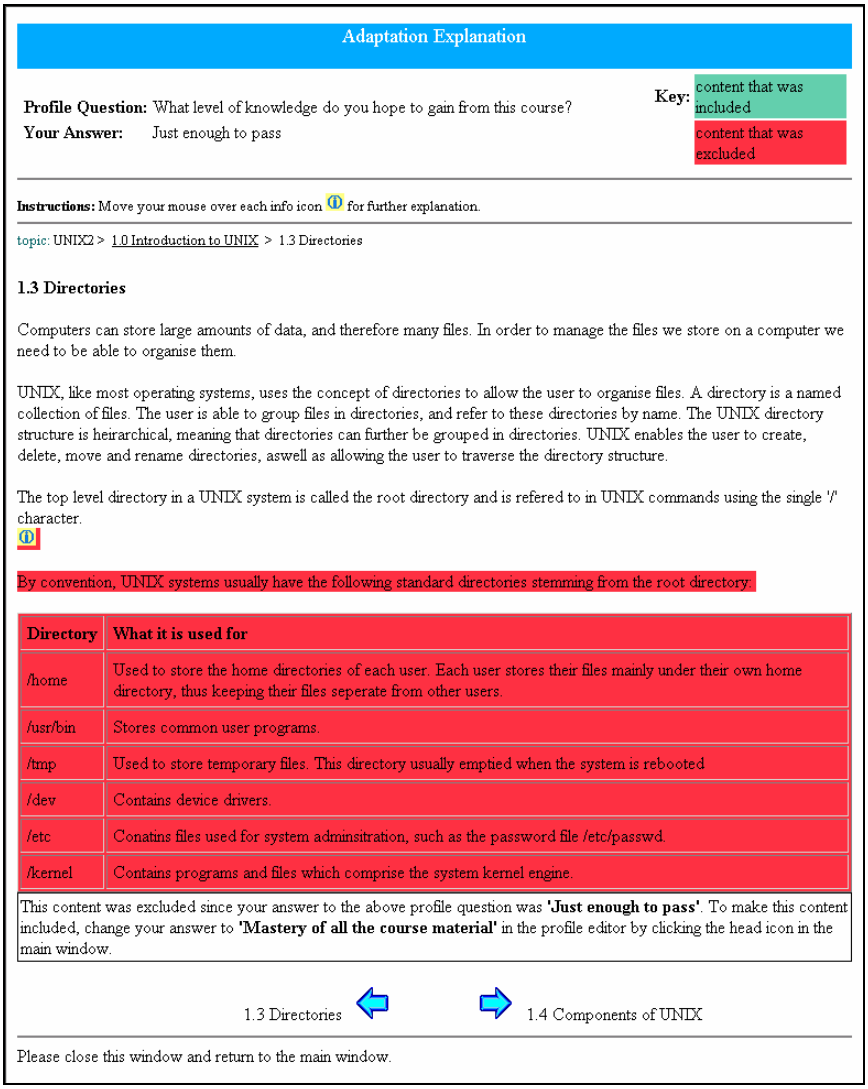
[hide explanation]

Figure 5.6 – Adaptation Summary in Tutor 3

The adaptation summary has several elements. It begins with information describing how to scrutinise the adaptation. Also, the text at the bottom of the adaptation summary reminds the user that they can change their profile settings by clicking the user profile icon in the top menu.

The core of the adaptation summary is the list of each user model attribute that caused adaptation of the page. This is expressed in terms of the profile questions so that it should be familiar. Against each question is the user's current answer. The third column displays a hyperlink labelled *show me*. For example, the last row in Figure 5.6 indicates that there has been adaptation because the student answered *Just enough to pass* for the profile question *What level of knowledge do you hope to gain from this course?* Clicking the *show me* hyperlink opens a separate browser window to display the Adaptation Explanation page in Figure 5.7. Using colour coded background colour for adapted content, it shows adaptation associated with the single profile question that the user selected with the *show me*

link. It is in a separate window so that the user can compare this annotated version of the page with the original version.



Adaptation Explanation

Profile Question: What level of knowledge do you hope to gain from this course?
Your Answer: Just enough to pass

Key: content that was included / content that was excluded

Instructions: Move your mouse over each info icon for further explanation.

topic: UNIX2 > 1.0 Introduction to UNIX > 1.3 Directories

1.3 Directories

Computers can store large amounts of data, and therefore many files. In order to manage the files we store on a computer we need to be able to organise them.

UNIX, like most operating systems, uses the concept of directories to allow the user to organise files. A directory is a named collection of files. The user is able to group files in directories, and refer to these directories by name. The UNIX directory structure is heirarchical, meaning that directories can further be grouped in directories. UNIX enables the user to create, delete, move and rename directories, aswell as allowing the user to traverse the directory structure.

The top level directory in a UNIX system is called the root directory and is refered to in UNIX commands using the single '/' character.

By convention, UNIX systems usually have the following standard directories stemming from the root directory:

Directory	What it is used for
/home	Used to store the home directories of each user. Each user stores their files mainly under their own home directory, thus keeping their files seperate from other users.
/usr/bin	Stores common user programs.
/tmp	Used to store temporary files. This directory usually emptied when the system is rebooted
/dev	Contains device drivers.
/etc	Conatins files used for system adminsitration, such as the password file /etc/passwd.
/kernel	Contains programs and files which comprise the system kernel engine.

This content was excluded since your answer to the above profile question was **'Just enough to pass'**. To make this content included, change your answer to **'Mastery of all the course material'** in the profile editor by clicking the head icon in the main window.

1.3 Directories 1.4 Components of UNIX

Please close this window and return to the main window.

Figure 5.7 – Example of an Adaptation Explanation page in Tutor 3.

(Accessed by clicking the show me link in Figure 4 for the profile question What level of knowledge do you hope to gain from this course?)

The colour choices follow a traffic light metaphor. This is explained to the user in the key at the top right of the adaptation explanation page. It shows the selected profile question, the students current answer and explains the colour coding. For example, in Figure 5.7, the fourth paragraph under the section titled *1.3 Directories* and the table below it are highlighted in red, indicating the content was adaptively excluded from the page. Moving the mouse over the *i* icon (just above the red paragraph) pops up the following text informing the user why the context was excluded, and how to change this adaptation:

This content was excluded since your answer to the above profile question was 'Just enough to pass'. To make this content included, change your answer to 'Mastery of all the course material' in the profile editor by clicking the head icon in the main window.

From the adaptation summary table on the main page Figure 5.6, the user can open a separate adaptation explanation page for each profile attribute that affected adaptation and thus compare how different profile questions affected the adaptation of the page.

5.3.3 Evaluation

In the evaluation of Tutor 2, some participants were told how to query the system for an adaptation explanation. They were instructed to click the hyperlink *How was this page adapted to me?* to view the adaptation explanation. As a result, it could not be determined whether participants would have known how to do this without help from the evaluator, and in turn whether users would be able to find the adaptation explanation in practice. In this evaluation, the participants were asked to investigate how a page had been adapted but they were not told how to do this.

For a detailed report on the evaluation of Tutor 3, refer to Czarkowski & Kay (2003a, 2003b & 2006). Each participant was provided with a worksheet, included in Appendix 8.3 (page 233). This described the learning goals, interests and background of a fictitious student, Fred. Participants were asked to assume the role of Fred and use Tutor 3 to work through an Introductory UNIX course. Participants were presented with one page of the worksheet at a time so they could not jump ahead. Participants we allowed to spend as much time as required. We observed the participants as they completed tasks to record their comments. In addition, the system logged all their interactions with the interface. The participants were not given any training or a demonstration of the system.

Each participant therefore had to set up their user profile with the same values and perform the same actions in the system to explain the adaptation. This allowed results of participants to be compared as they were expected to perform the same system actions and write the same answers in the worksheet, unlike the evaluation of Tutor 2. Following Nielsen (1994), we selected five participants for the evaluation. They have different backgrounds and varying degrees of computer literacy: one was a secondary school student, two were third year computer science degree students and there were two adult participants with basic computer literacy skills.

The tasks of the worksheet were designed so that each basic issue was explored in three subtasks. This provided internal consistency checks, an important concern since there are degrees of understanding and we wanted insight into just how well each participant was able to scrutinise the adaptation. Extracted questions from the worksheet are shown in Table 5.2. These reflect our core evaluation goals as described at the beginning of this section.

Table 5.2 – A sample of questions in a worksheet completed by participants in the evaluation of Tutor3.

Part	Concept	Questions presented in worksheet
1	Understanding the purpose of profile questions	• Will Tutor use Fred's profile settings? If so, what will it use Fred's profile settings for? • Where does Tutor get information about Fred to use to perform adaptation of the content to suit him? • Will Fred be able to influence or control the way Tutor adapts content to him? If so how?
2	Ability to determine what was adapted	• What would you do in the system to find out whether Tutor adapted any material on the Teacher's Instructions page to you? • Did your answer to the profile question *What is your main objective?* Have any effect on the contents of the Teacher's Instructions page? How do you know this? • Was any content specifically excluded because of your answer to the profile question *What is your main objective?* If so, what was the content?
3	Demonstrating control over adaptation by changing answers to profile questions	• Now consider what the first sentence on the page would read had you answered the profile question *What is your main objective* with *Revise the material*. Without changing your answer just yet, write out how you think the sentence would read. *By the end of this course you will have …* • Explain what actions you would have to perform in the system to change your answer to the profile question *What is your main objective?* • Change your answer for the profile question *What is your main objective* to *Revise the material*.

The first task was to complete the profile of Fred by answering questions on the profile page. The worksheet did not specify the answers for each question but did describe Fred's learning goals, interests and background. All participants correctly answered these profiles questions consistently with the information we provided about Fred which meant that the pages displayed to them in the experiment had the same adaptations.

Each participant was then asked questions from the first block in Table 5.2. In their answers to the first two questions, all participants stated their profile answers would influence the way material would be presented to them. However, participants had difficulty with the question *Will Fred be able to influence or control the way Tutor adapts content to him? If so how?* Most participants indicated that although they believed they could initially influence the system through their profile answers, they expected that once they started working through the course they would no longer be able to influence the system or change their profile. All seemed to believe there were factors they could not control influencing adaptation. Note that this is despite the help text on the profile page, as described in the previous section.

The second part of the worksheet asked each participant to navigate to a specific page. Then, participants were asked to determine whether any of the content on that page was adapted to Fred, based on the answer to the profile question *What is your main objective?* The participants had to indicate any adapted content and whether it was included or excluded. To perform this task, the user had to notice and click the hyperlink *How was this page adapted to you?* We expected participants would find this exercise straightforward. However, this was quite challenging for most participants. Only two of the five participants completed this task as expected. We noted these participants examined the interface carefully. The three other participants did not see the hyperlink *How was this page adapted to you?*

The third part of the worksheet involved what-if experiments. Participants were asked to guess the content of a specific paragraph, assuming Fred were to change his answer for the profile question *What is your main objective?* To *Revise the material*. Answering this question should not have required any guesswork since the answer provided by the Adaptation Explanation. We expected participants would examine the adaptation explanation which shows how the content would be affected based on a change to the profile. Again only two participants (same participants as above) were able to complete this task.

Overall, only two out of five participants were able to effectively use all the scrutinisation tools.

5.3.4 Conclusion

Our evaluation identified some important and interesting outcomes that are important for our goal of supporting scrutability and control over adaptation.

(1) The participants in our study were comfortable our simple user models used to capture information to achieve personalisation

Our users did understand the concept of a user model. They appreciated that the system would store information about them and in return provide personalised material.

(2) The participants in our study understood that they had input to the personalisation process, but needed convincing that they could control it as well

Having filled in their initial profile, all participants could appreciate that their user model would influence the personalisation of material. However, three out of five of our users needed time to work out that they could control the personalisation. They appeared to believe that once they filled in their profile, that would be the end of their input. It seems that our users, perhaps due to previous experiences, assumed they did not have any control over the adaptation. One user expressed surprise at the extent to which they could control the adaptation. Notably, all two users learnt they could control adaptation and they achieved this by reading the instructional text presented by the system; no intervention was required.

(3) Building an interface for scrutability support is difficult.

A key concern, highlighted by the experiment, was that users had difficulty finding the scrutability tools when needed. Most participants required help to find the link *How was this page adapted to you?* Although we had provided instructional text explaining how to access the adaptation explanation and profile page, users who skimmed over the interface tended to ignore this. They then had difficulty completing the evaluation. Also, although Tutor 3 has online help, none of our users accessed it.

Overall, it seemed that the notion of adaptivity was more familiar to our users than in the Tutor 2 evaluation. This is consistent with the growing use of adaptivity in the time between when these two evaluations were conducted. However, the idea that they could control the adaptation was new.

On the basis of the Tutor 3 evaluation, if we imagine that these users had been using Tutor 3 in an authentic learning context and they had cause to wonder about the adaptivity, it is unclear whether they would have thought to try to scrutinise the adaptivity. Even if they had found the relevant links, it is seems that some would not have been able to work out, unaided, exactly what was adapted and how.

5.4 Cell-Tutor

Following the evaluation of Tutor 3, our goal was to refine the interface to the scrutinisation tools so that users could find them more easily. Collaboratively, a new version of the system was developed and evaluated. The system was not titled so we refer to it as the Cell-Tutor version (Czarkowski, Kay & Potts; 2005a, 2005b).

5.4.1 Design Goals

The idea, proposed by the research group supervisor Judy Kay, was to provide a blatant, always-present reminder that adaptation is being performed in a key position in the interface. This would surely capture the user's attention and help them find the scrutinisation tools.

5.4.2 System Overview

The system was built by Serena Potts using a local lightweight but highly adaptable web framework called Cellerator (Kummerfeld & Lauder, 2003). Taking the categorisation by Bicking (2003), it is a script-based framework in conjunction with Personislite (Kay, Kummerfeld & Lauder, 2002), which provides scrutable user modelling. Following the terminology of Brusilovsky (2000), the system provided forms of both adaptive presentation and adaptive navigation. It also drew on many elements of the Tutor 3 system. Essentially, it took the rather radical approach of always presenting the student model along with the adapted content on every page.

A typical page is shown in Figure 5.8. The 'Personalisation' and 'User Model' cells are present on every page, allowing the user to see the values of their user profile at any time. As in Tutor 3, if a page has been adapted, '*How was this page adapted to you*' is displayed at the end of the coursemap or lesson page. Alternatively to indicate that a page has not been adapted, '*This page has not been adapted*' is displayed at the end of the page. By default, the page displayed for a user has been adapted based on their user model. They cannot see what has been adapted or why. By clicking on link '*How was this page adapted to you?*', the same page will be displayed, this time highlighting sections that were included based on the user's user model in yellow and sections that were excluded in green, as in Tutor 3.

Figure 5.9 shows what happens when the link *How was this page adapted to you?* Is clicked for typical page. To see why an individual section has been included or excluded, the mouse is held over the section in question, and a caption will pop up to indicate the reason. The figure shows an example of the use of the mouse-over to show why content was included where the caption is displayed: '*This was included because your level was: basic*'. Similarly, if the mouse was held over an excluded piece of text, a mouse-over explaining the reason for the exclusion would be presented. This is similar to Tutor 3 except that here, moving the mouse over any part of the highlighted portion of reveals the explanation. In Tutor 3 the user had to aim for the '*i*' icon to reveal the information, which was a smaller target.

To change their profile, the user can always click '*Change your profile'* in the Personalisation cell to display their profile page, as in Tutor 3.

In summary, the key difference from Tutor 3 the Personalisation cell. It is displayed on every page, given a significant amount of key page real-estate where the user is likely to see it and be reminded that personalisation is taking place.

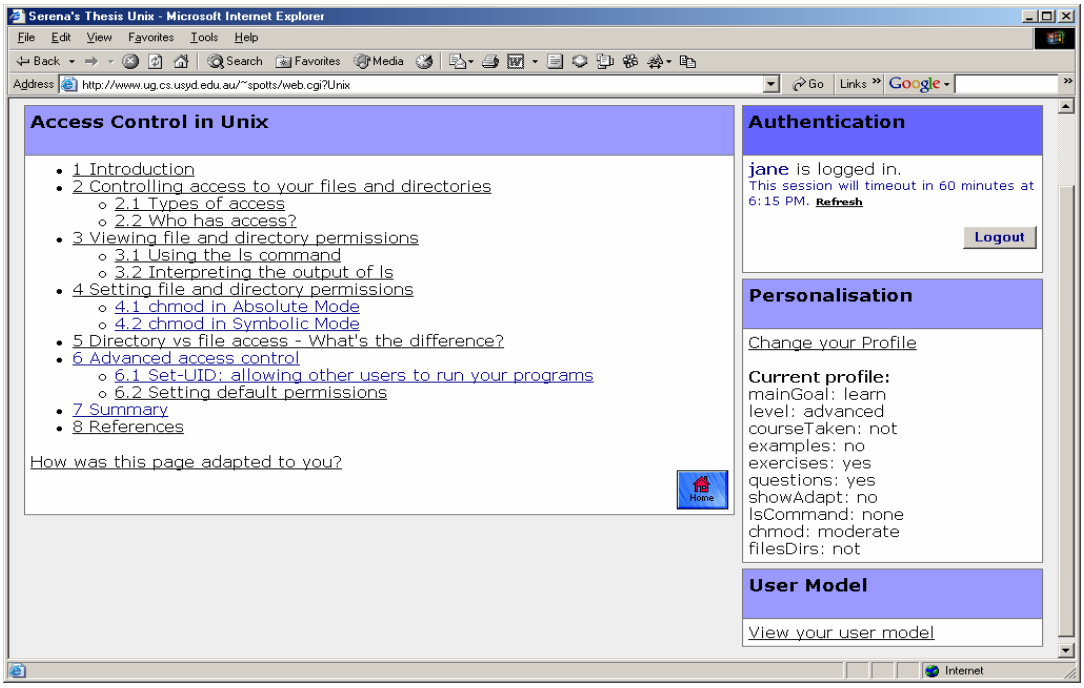


Figure 5.8 – Typical page in Cell-Tutor version. The right had side of the page shows the user model.

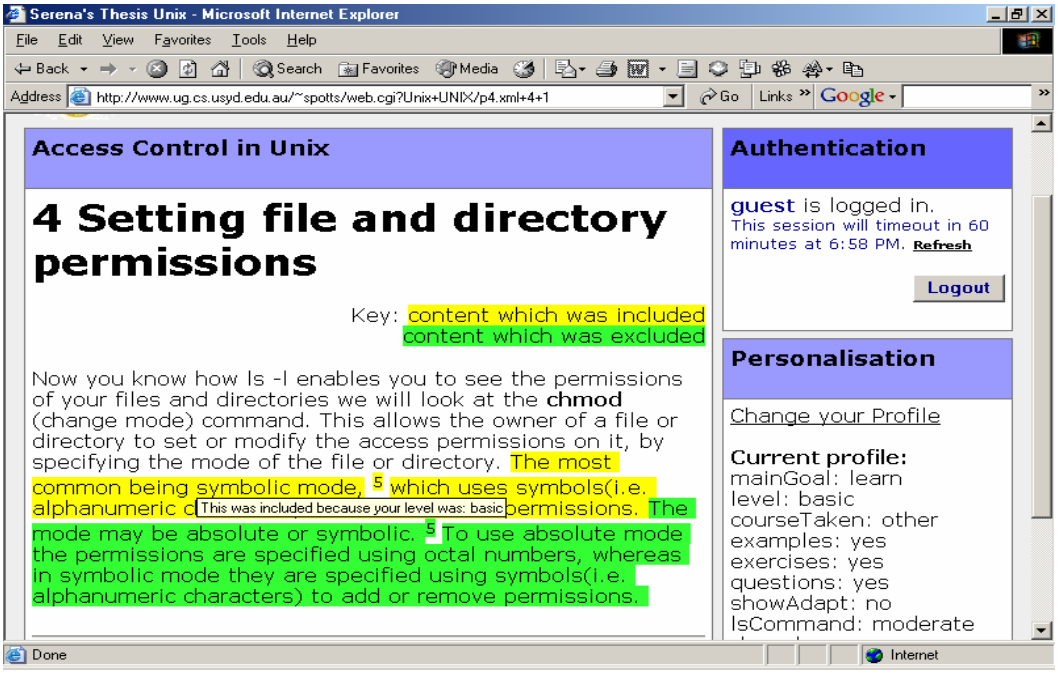


Figure 5.9 – Mouse-over showing content that was included indicated by the caption: 'This was included because your level was: basic'.

5.4.3 Evaluation

The evaluation approach was similar to that of Tutor 3. The evaluation was carried out by Judy Kay and Serena Potts. A detailed report is published in Czarkowski, Kay & Potts (2005a).

The evaluation was undertaken in two groups. In the first group, Participants 1-5 were asked to answer the initial questionnaire on the profile page as if they were a single fictitious user, Fred, as in the Tutor 3 evaluation. By contrast, Participants 6-9, in the second group, answered for themselves. In the first group, all users had similar experiences and so we could frame tasks and questions accordingly. The second group sacrificed this comparability and consistency but complements the evaluation by providing insights into the experiences of users who do not have the potential cognitive burden of remembering the profile elements for Fred. The study explored both a consistent but inauthentic user model (Group 1) and the authentic Group 2 user model, where different participants, giving different answers to the user model questionnaire, would experience different adaptivity.

Having filled in their profile, participants were then instructed to continue working through to the end of the tutorial, then answering questions relating to their understanding of how content is adapted by the system and why it was included or excluded based on their user profile. When they had completed the tutorial, participants in Group 1 answered questions which assessed whether they could predict how content would change if they altered their profile. We expected them to use mouseovers to see what content was included or excluded.

An error in the personalisation was purposely included in the adaptive content. We reasoned that this is just the sort of situation in which it would be natural for a user to want to scrutinise the personalisation to determine why the system appeared to behaving unexpectedly. We wanted to see if participants identified that the system had adapted something incorrectly based on their profile. The errors involved the presentation of exercises on every content page even though the answer to the profile question about trying relevant exercises was answered as ‘No’ initially.

All participants understood the role of the profile and that changing the profile alters adaptation. Most understood coloured sections showed content that was included or excluded by personalisation.

All Group 1 participants understood that their user profile would cause the adaptation of content within the system and were able to effectively change their user profile. Most were able to view the adaptation though of those who could, the majority experienced difficulties utilising the mouse-over function provided to see the reason for the adaptation.

We asked affective questions of participants in Group 2. All but one participant favoured the idea of adaptation. All gave positive feedback on the system design with its clear layout, good navigational tools and helpful instructions and explanations as appropriate.

The ‘bug’ in the system was not uncovered in either stage, even when participants were informed in Group 2 that there was a bug. This may have been because the learners were very absorbed in the learning tasks: the initial questionnaire indicated that the participants did not know the content of the tutorial initially and all reported satisfaction in the quality of the learning experience. Participants in Group 2 appeared to be more motivated than those in Group 1, which may have contributed to their understanding of the system, as they were willing to learn and may have been paying more attention.

5.4.4 Conclusion

The significant result was that all participants in the evaluation of this new system understood that a profile change alters adaptation. We attributed this result to the “Personalisation” cell on every page, which informed users that personalisation was occurring on every page, showed them their current profile (user model) attributes and provided a link to the profile page to change them. Compared to the Tutor 3 evaluation, users understood more deeply the relationship between the profile and personalisation.

However, users still had difficulty determining what caused specific adaptations and in particular using the mouse-over tool.

5.5 SASY 1

With the evaluation of the Cell-Tutor, we concluded that displaying the user profile in a significant portion of the screen on each page helped users understand that personalisation was based on their profile and that they could change personalisation by changing their profile. The evaluation showed that the concepts of personalisation and user modelling were not as foreign to users as we had thought in our evaluations of the Tutor systems. Partly, this was because the use of personalisation had become more common on internet sites. Partly, the Cell-Tutor evaluation showed a clear and intuitive interface could help bridge the gap to introduce users to scrutinisation.

5.5.1 Design Goals

There were still interface challenges to resolve. Firstly, users still had difficulty determining what personalisation occurred on a page. Observations of users during the Tutor2, Tutor3 and Cell-Tutor evaluations showed that users either did not see, or did not think to click the *How was this page adapted to you?* Link at the bottom of a personalised page. Secondly, once they did click it, they had difficulty using the mouse-over explanations. Cell-Tutor made the user model (profile) more visible than in previous systems and this made the user model scrutinisation tool more noticeable and accessible to participants. With SASY, we aimed to extrapolate this idea to increase the visibility to the other scrutinisation tools.

In addition, we wanted to explore the role of scrutinisation with more complex personalisation that is representative of real-world adaptive systems. We had some success with the simple user modelling approach in the Cell-Tutor system, where personalisation was based on the user’s profile that was only changed directly by the user. The next logical step was to explore tools to support scrutinisation where

the user model was also updated by the system based on observations and inferences about the user. This added another level of complexity to the personalisation and this, in turn, increased the difficulty in developing scrutinisation tools to support this.

5.5.2 System Overview

In SASY 1, we changed the interface and significantly enhanced the underlying user modelling and personalisation capabilities of the system. This section describes the enhancements in detail. Where SASY 1 has the same functionality as SASY 4, the user is referred to read chapter 3 rather than duplicating the information here.

User Modelling & Personalisation Enhancements

The SASY 1 user modelling and personalisation functionality is the same as for SASY 4, described in Chapter 3. In summary, the following enhancements were implemented in SASY 1. Where functionality is the same across all SASY systems, we refer to the system as SASY.

Direct and Inferred user model attributes:

In previous systems, all user model attributes had to be set by the user and did not change unless the user changed them. The systems were adaptable rather than adaptive. In SASY, user model attributes can either be set by the user or inferred by the system. The first time a user accesses a SASY topic, they are still presented with an initial profile questionnaire to elicit some characteristics about them. Other attributes can default to constant values or be calculated, based on arbitrary logic expressions involving other user model attributes. If the user returns to their profile at a later stage they would see the values that have been inferred and may change these values in the same way they change values for attributes they volunteered in the initial profile.

Dynamic updates to user model

SASY may update attributes in the user model following an event such as:

- User access to a page.
- User mouse click.
- User selecting an answer to a quiz question.

User model Evidence

Since user model attributes could be updated by the system, the user model was extended to include a history of events that caused an attribute to be changed. The evidence is displayed to the user through the Evidence Tool (discussed below) to explain why the profile has been changed.

Personalisation Enhancements

In SASY, the condition specified for whether adaptive content should be shown to a user can be expressed as an arbitrary, python code expression involving other user model attributes. This provided flexibility as the expression gave flexibility to the author. If a complex expression was provided, the

author may also provide a text description that is shown to the user explaining why the content was added or removed. Otherwise, the system generates an explanation based on the condition.

Scrutinisation Tools

A typical personalised page in SASY 1 is shown in Figure 5.10. The right hand side of the page is devoted to scrutinisation. As in the Cell-Tutor system, a significant amount of the main portion of the screen is devoted to personalisation in order to be in plain sight. We called this the Personalisation panel.

Underneath the "Personalisation" heading, are the Removed and Included items indicators. These specify how many content items have been removed and included by the personalisation on that page. The indicators are hyperlinks that invoke the Highlight Tool, which applies a background colour to content that has been personalised on the page. Below this is the hyperlink *Click to highlight removed / included items on this page*. Clicking this link also invokes the Highlight Tool.

Below that is a list of the profile attributes that had affected personalisation on the current page, called Evidence Links. This is similar to the Personalisation cell in the Cell-Tutor system (see Figure 5.8) but with a number of enhancements. Firstly, SASY 1 displays human readable descriptions of profile attributes (e.g. *You want lots of practice quizzes*) rather than variable name/value pairs (e.g. *exercises: yes*). Secondly, only attributes that affected personalisation on the current page are displayed. This means the list will vary from page to page and therefore is more likely to attract the user's attention. This may also be very important if there is a large user model but only a few components play a role in personalisation on each page. SASY 1 ensures such pages have a short, more readily understood, list of personalisation details. Thirdly, clicking the Evidence Links invokes the Evidence Tool that pops up a small window listing the history of events that caused the profile attribute to have its current value. This is the same as the Evidence Tool in SASY 4, in Chapter 3.

The Profile Tool can be accessed either from the menu at the top of the page (Figure 5.10) or from the Evidence Tool. The Profile Tool is the same as in SASY 4, see chapter 3.

An example of the Highlight Tool in SASY 1 is shown in Figure 5.11 when it is invoked from a typical page (Figure 5.10). Instruction text is included at the top of the page when the Highlight Tool is invoked *Hold your mouse over a coloured section to see the reason for the personalisation*. The colours used for colour coding differ from Tutor 3 and the Cell-Tutor version. We chose more subtle colours. We believed users would be most keen to see the removed content and thus this should attract the users attention more than content that had been included by personalisation. Here we used a bright yellow for removed content (so that is stood out) and a dull grey for included content (so that it faded in to the background). Each personalised content item was surrounded by a thin black border. This allowed the user to see nested adaptive content, as in Figure 5.11.

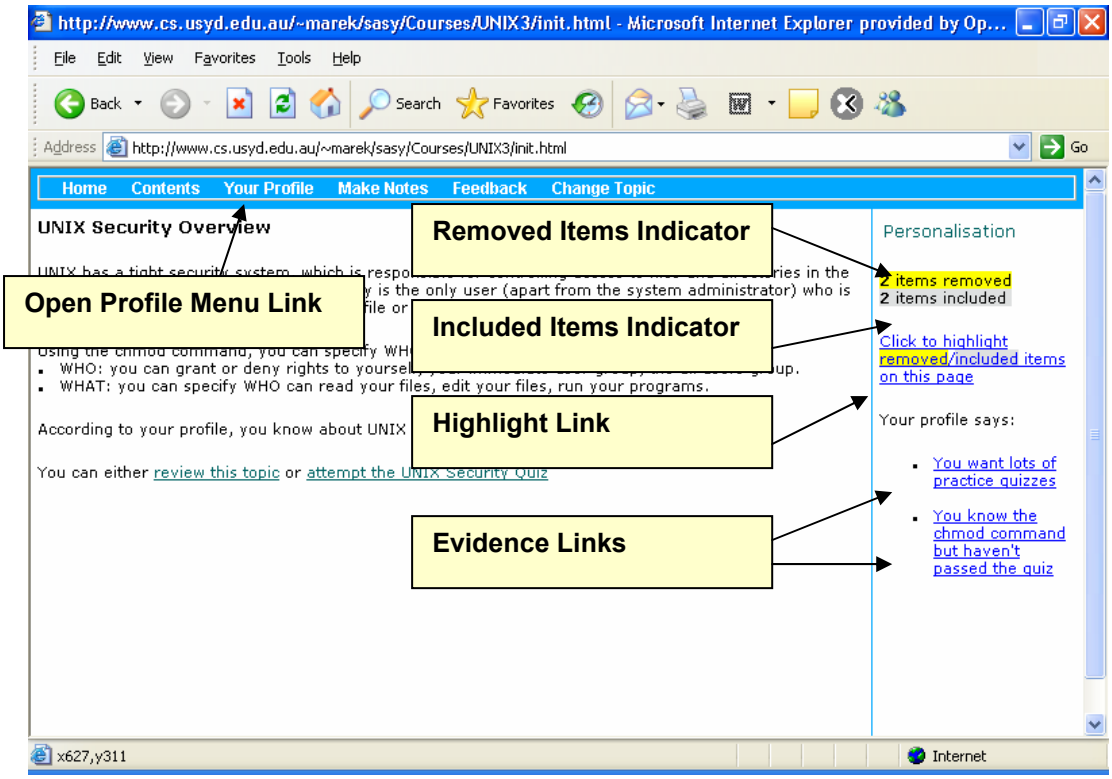


Figure 5.10 – Typical personalised page in SASY 1 with scrutinisation tools pointed out.

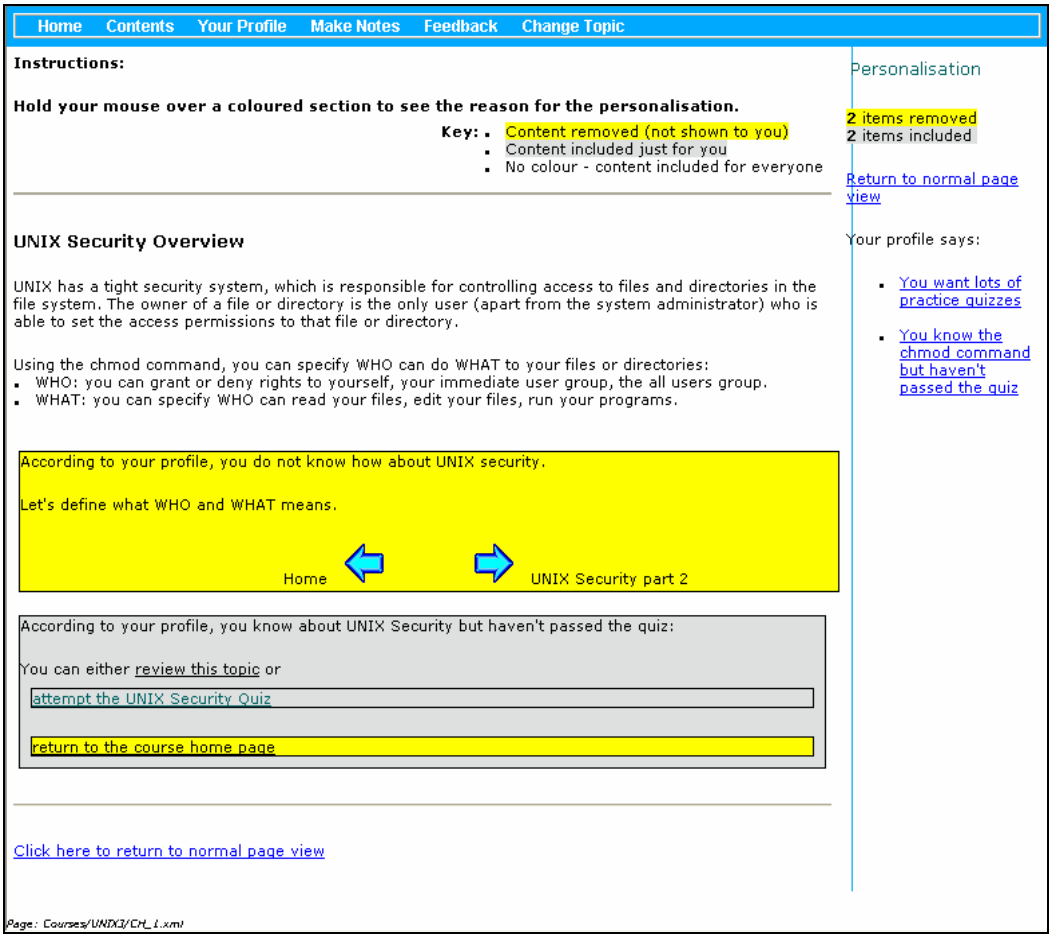


Figure 5.11 – Shows example of Highlight Tool invoked from page in Figure 5.10.

Moving the mouse over a highlighted area on the page pops up a message explaining why the content has been removed or included by personalisation. An example is shown in Figure 5.12, where the message is popped up *Not shown to you because your profile says you know the chmod command but haven't passed the quiz*.

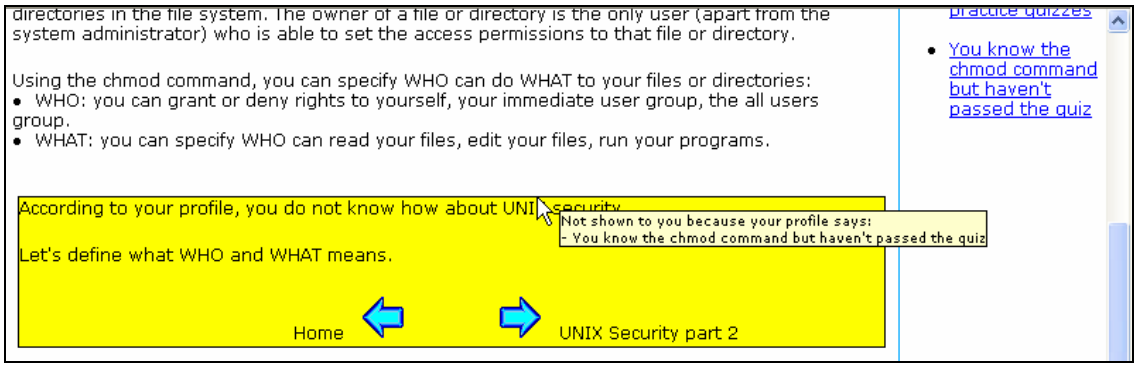


Figure 5.12 – Shows example of Mouse over explanation of Highlight Tool.

In summary, there were two key differences between SASY and the Cell-Tutor version.

Firstly, we introduced several elements to the right hand side of the page to summarise the personalisation and allow the user to invoke the Highlight tool to scrutinise in more detail. Based on the Cell-Tutor evaluation outcomes, we expected users would quickly notice these new elements and this would solve the problem of users being unable to find the link *How was this page adapted to you?* At the bottom of a page and hence could not tell how the page was personalised.

In addition, we greatly expanded the user modelling and personalisation capabilities of the system. SASY 1 used data provided by the user in their initial profile questionnaire and also made inferences about them as they used the system, so that the user model was dynamic rather than static. To help users scrutinise this more complex personalisation we added the Evidence Tool that provided the history of events, or a list of evidence, that caused SASY to update a profile attribute: in a sense, to justify its actions to the user.

5.5.3 Evaluation

We believed the scrutinisation toolkit in SASY 1 was mature enough for a detailed evaluation to determine whether the users would use the tools in an authentic environment and whether the tools were effective and easy to use. Chapter 6 reports the two evaluation studies for SASY. Based on the studies we further refined the tools to the form they have in SASY 4.

5.6 Summary

This chapter described the progressive development of the scrutinisation tools. The goal was to design and develop tools to support various forms of scrutinisation as described in Chapter 1. Furthermore, the tools should not require any training and users should be able to use the tools effectively without prior experience.

Developing an interface to support the process of scrutinisation is difficult. One would we expect users will not scrutinise often as it is a distraction from their main task. One might suggest the scrutinisation tools should be hidden away in the interface unless they are needed. But when users need to scrutinise, perhaps to correct a system misconception, they need to easily find and access the scrutinisation tools. Ideally, the tools should not require any training and users should be able to them effectively without prior experience or if have not used them for a long time.

We started with a very simple approach in Tutor (versions 1 through 3) where the user model was built solely by the user's answers to an initial questionnaire. The user model was used to adapt web pages that taught UNIX concepts. The scrutinisation tools were tucked away at the bottom of every adapted page. Although the personalisation was very basic, this made a logical starting point for exploring ways to scrutinise to adaptation. Surprisingly, users were not able to use the scrutinisation tools effectively and were overwhelmed by the number of explanations the tools provided.

We refined the user interface in Tutor 2 and Tutor 3 to provide several levels of detail for scrutinisation. The tools first provided a summary of the personalisation then allowed the user to delve into more detail should they wish to do so. The evaluation of these systems showed that users in our studies:

- were comfortable with simple user models used to capture information to achieve personalisation;
- understood that they had input to the personalisation process, but needed convincing that they could control it as well;
- were able to use the scrutinisation tools after some practice but could not always figure out how to access them when needed.

Next we tried a more radical approach, where the simple user model was always present and visible on every page (Cell-Tutor version). Evaluation showed an increased awareness of the user model and its role in the personalisation process.

This inspired us to make other scrutinisation tools more visible to the user and explore tools to support scrutinisation where the user model was also updated by the system based on observations and inferences about the user. This added another level of complexity to the personalisation and in turn increased the difficulty in developing scrutinisation tools to support this. We developed SASY to

explore the role of scrutinisation with more complex personalisation that is representative of real-world adaptive systems. Chapter 6 reports the evaluation of SASY in detail, and describes the evolution from SASY 1 to SASY 4, the final version.

Chapter 7 presents the conclusion, including guidelines for a developing tools to scrutinise an adaptive hypertext based on the research discussed in this chapter and evaluation in the next chapter.

Chapter 6 Evaluation

One central contribution of SASY is the set of *scrutinisation tools* that allow the user to scrutinise how an adaptive hypertext page is personalised to them. Specifically, the tools allow the user to scrutinise in the following ways:

- Understand what aspects of the hypertext were personalised. That is, what parts of the content were added to and/or removed to create the user's personalised view of the page.
- Understand what aspects of the user model caused each personalisation.
- Visualise the implications of changing their user model, in terms of how it affects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without actually making any changes.
- View and understand evidence of how user model attributes were set, whether set directly by the user or inferred by the system.
- Change aspects of their user model to control the personalisation.

The design and development of SASY 1 was informed by our previous work, where we developed and evaluated the series of systems: Tutor 1 (evaluated quantitatively); Tutor 2; Tutor 3; and Cell-Tutor (evaluated qualitatively). We had also performed a cognitive walkthrough of SASY 1 and informally observed one person using the system, obtaining their feedback. Therefore, we believed SASY 1 was ready for larger scale evaluation.

The following were the evaluation goals.

Goal 1 - Assess if users can work out how to use the scrutinisation tools in the ways listed above without training.

As scrutability is a new concept and the scrutability tools are a new form of software, the purpose of evaluation was to determine whether users were able to use the tools to scrutinise effectively in the ways listed above. They should be able to determine how to use the tools by themselves, without training. This is a requirement because we know from our previous work with scrutinisation (Czarkowski & Kay 2000), that people will scrutinise infrequently and they need to be able to work out how to use the tools when they chose to scrutinise.

Goal 2 – Given access to SASY's scrutinisation tools in an authentic adaptive hypertext application, determine what are usage patterns of the scrutinisation tools.

Essentially, we would like to know whether users scrutinise, even though they are not required to do so, how much they scrutinise and in what ways they scrutinise in an authentic environment. Such an environment has the following properties:

- The user's main purpose for using the adaptive hypertext system is something other than scrutinisation. For example, a student may be using a personalised teaching system to learn about UNIX File System security. Here the user's focus is to learn. The user may scrutinise the personalisation at their will, but they are not required to do so.
- The user is not monitored obtrusively.
- The user is not trained to use the scrutinisation tools. However, the user may access online help if they are interested in learning details of the tools.

Ideally, the evaluation of Goal 1 would also take place in an authentic environment. However, there are several difficulties in doing so. We know from our earlier work (Czarkowski & Kay, 2000, 2003) that some users will be very interested in scrutinising while others might not scrutinise at all. Users will scrutinise to different degrees, at different times and in different ways. We cannot determine when a user will scrutinise but we predict it is when they are surprised or irritated by the personalisation, or just curious about how it works or whether it can be improved to better suit their current goal or activity. This makes it impractical to evaluate the usability or effectiveness of the tools by observing users in an authentic environment, rather than a laboratory environment.

Goal 3 – Obtain user feedback about SASY's scrutinisation tools and on the notion of scrutability and control of personalisation.

The literature highlights the importance of scrutinisation and user studies with personalised systems (Alpert et al. 2003; Ackerman et al., 1999) report that the ability to scrutinise personalisation is a highly desired feature of adaptive systems. As SASY provides this feature, we wanted to capture feedback from users about their experience using the tools, their desire for, and the perceived importance of, the ability to scrutinise a hypertext that has been personalised to them.

Goal 4 – Evaluate across different subject matter domains and styles of adaptive hypertext system.

SASY has been designed to provide a shell for creating scrutably adaptive hypertexts. It should be able to support adaptive hypertexts across varied domains: for example, it should be able to support creation of an educational adaptive hypertext and equally, a quite different adaptive hypertext such as ecommerce, or a recommendation system.

We would like to evaluate SASY, with its support for scrutinising the personalisation, across different domains. In addition, we wish to evaluate across different styles of adaptive hypertext system as this is a goal of SASY's scrutable adaptive hypertext framework. We study SASY when used as an adaptive hypertext teaching system and as an adaptive hypertext recommendation system.

6.1 Evaluation Design Overview[12]

To achieve the above evaluation goals, we divided the evaluation in to two experiments. Table 6.1 shows how the four evaluation goals are divided into the two experiments.

Table 6.1 – Grid showing the evaluation goals of Experiments 1 and 2.

	Experiment 1	Experiment 2
Goal 1	YES (Limited)	YES
Goal 2	YES	NO
Goal 3	YES	YES
Goal 4	YES	

The first experiment is an unobtrusive field study, where students used SASY 1 to learn about UNIX Security in an authentic environment. In this study, SASY 1 was used as an adaptive hypertext teaching system. The purpose of this experiment, as in Goal 2, was to gain evidence about the way users scrutinise in an authentic adaptive hypertext application, where their main goal for using the system is to learn about UNIX Security. Actions of 105 student participants were logged, in the background, in order to allow us to analyse usage patterns. At the end of their SASY session, students completed an online questionnaire that captured their qualitative feedback (Goal 3 and Goal 1 to a limited degree).

A second experiment reports a laboratory-based, qualitative study where we asked users to complete tasks around a Personalised TV Guide. Here SASY 1 was utilised as an adaptive hypertext recommender system that recommended a personalised television program viewing schedule to users based on their specific interests. The purpose of this experiment, as in Goal 1, was to assess the effectiveness and usability of the scrutinisation tools, this time in a completely different domain. We explored four system variants (SASY 1, SASY 2, SASY 3 and SASY 4) to refine the scrutinisation tool interface. Users were not directly asked to use the scrutinisation tools but had to realise how to access and use the tools in order to complete the tasks. At the end of their SASY session, users completed an online questionnaire, providing feedback. Additional feedback was obtained through discussion with participants following the evaluation (Goals 1 & 3).

Together, the experiments evaluate of SASY across different subject matter domains and adaptive hypertext system styles (an adaptive hypertext teaching system in Experiment 1, an adaptive hypertext recommender system in Experiment 2), as discussed in Goal 4.

[12] This evaluation design was informed by an initial ideas published in Czarkowski (2005c).

Though we would have preferred to carry out Experiment 2 (detailed laboratory study) before Experiment 1 (field study), this was not possible for a couple of reasons. Firstly, it was due to timing constraints. Our target participant audience for Experiment 1 were third year computer science students at the University of Sydney enrolled in a particular subject that required them to learn about UNIX Security. This class is only taught during the first semester of each year. Since the evaluation materials were ready, we needed to conduct the experiment at the time they needed to learn this material. Secondly, the design of SASY 1 was informed by the development and evaluation of several previous systems, so initial problems with user interface had already been dealt with. In the evaluation of Cell-Tutor, the predecessor to SASY, the majority of participants could identify that their profile caused the adaptation, were able to see what had been adapted to them; understood why it had been adapted; could change their user model, hence controlling the adaptation. Cell-Tutor made the user model more visible than in previous systems and this made the user model scrutinisation tool more noticeable and accessible to participants. With SASY, we extrapolated the idea to increase the visibility to the other scrutinisation tools.

Overall, the participants in our experiments represented a well educated, computer literate group. To some degree, understanding the information provided by the scrutinisation tools requires the ability to think logically. People with a moderate level of computer literacy and those trained to think logically, such as computer science students, probably have a better chance of being able to use the scrutinisation tools compared to people without such a background. As such, it is reasonable to first evaluate with a sample group that has a good chance of understanding the concept of scrutinisation. Then having found the users could scrutinise, broaden the sample group to better represent wider and more diverse community.

We now report experiments 1 and 2 in detail.

6.2 Experiment 1 – Field Study

6.2.1 Design

The purpose of this experiment was to analyse usage patterns of SASY's scrutinisation tools in an authentic adaptive hypertext application (**Goal 2**) and to obtain user feedback about the tools (**Goal 3**). We created a SASY course that personalised and presented UNIX File System Security concepts. This topic was part of the curriculum of a large number of students who genuinely needed to learn this material. The course consisted of a pre-test, the presentation of UNIX concepts, quiz questions followed by a post-test. In the background, SASY logged user actions to a log file monitoring usage of scrutinisation tools. The pre and post scores were used to measure user knowledge about UNIX File System security. The post-test was also used to obtain user feedback through a qualitative survey. Following the user sessions with SASY, we invited users, by email, to provide additional feedback and discuss their experience using SASY's scrutinisation tools.

Our hypothesis was that users will scrutinise and control personalisation if they perceive value from doing so. For example, if the personalisation is blaringly inappropriate, this should drive a user to scrutinise and correct it. Since it is possible that users might not scrutinise at all, to motivate them to do so, we set up their initial user model so that SASY's personalisation would produce unexpected, perhaps surprising results. This approach was inspired by the well known article "Oh no! My TiVo thinks I'm gay"by Zaslow (2002), where a user of a television recorder is concerned the device believes he is gay, but the user has no easy means to change its perception. Table 6.2 lists the user model defaults and the types of scrutinisation each was designed to motivate. We believed the jokes and hints, advanced concepts and tedious quiz questions that we added to the material would be seen as annoying by some students. At the same time, this represents a personalisation design decision that an author might plausibly make. We expected these user model defaults would motivate users to scrutinise how a page was personalised, scrutinise their user model, scrutinise evidence of how a user model attribute was set, and/or change their user model.

Table 6.2 – User model defaults and the types of scrutinisation they were designed to motivate.

	User model defaults designed to motivate user to scrutinise	**Expected Scrutinisation**
A	Add jokes and hints to the course material.	Remove jokes and hints from the material.
B	Teach advanced concepts of all topics rather than just essential content.	Remove advanced concepts from the material.
C	Require users to answer many quiz questions to test their knowledge of the material.	Only show a small number of quiz questions included in the material.
D	Claim the user had no knowledge about a UNIX security concept unless they correctly answered every question about that concept in the pre-test. This meant some students were presented with content as well as quiz questions about concepts they already knew quite well.	Reflect the user had some knowledge about a concept though they did not answer every question about that concept correctly in the pre-test.

We analysed the logged actions to determine which scrutinisation tools were used, how much they were used, on which pages and at what time during a session. We expected the logged actions to show evidence of the forms of scrutinisation listed above as well as other natural forms, for example, users scrutinising or changing other aspects of their user model.

6.2.2 Participants

Third year computer science students participating in the User Interfaces and Design Programming class at the University of Sydney were asked by their lecturer to complete the UNIX Security course in SASY as part of their homework assignment for a particular week of the semester, 11th April – 15th April 2005. They were asked to complete the SASY course in their own time, from their chosen

location and present a print-out of their completed post-test to their tutor as part of their homework preparation.

Having removed all actions of lecturers and supervisors from the logs, 105 students used the UNIX Security course, 82 completed the post-test.

6.2.3 Materials

Users typically completed the pre-test, then visited the pages of the course content and completed quiz questions for three modules: UNIX Shell, File System and Chmod command, then completed the post-test. Users were free to use SASY over multiple sessions. This section first describes the design of the pre and post tests and then the course content.

Pre-Test and Post-Test

The pre-test consisted of ten multiple choice questions about the UNIX Shell, File System and Chmod command. Unlike regular SASY quiz questions, selecting an answer did not provide feedback immediately. Students were only given feedback as to what they answered correctly once they answered all the quiz questions. The system allowed users to attempt the pre-test only once. An error was displayed if student attempted to access or submit a test for the second time.

The post-test followed the same format and had similar questions to the pre-test and in addition included qualitative questions about SASY's scrutinisation tools, as shown in Table 6.3. The system allowed users to attempt the post-test only once. The purpose of Questions 11 and 12 was to determine whether users knew they were able to change the personalisation to remove hints and jokes from their personalised view. Questions 13 established whether users believed they knew how to inspect the personalisation to see what items were added or removed from the personalised view. Questions 14 established whether users considered they could alter the personalisation by changing their user model. Question 15 asked users whether users they believed it was useful to be able to inspect and control the personalisation. Question 16 asked users who did not inspect the personalisation, why they did not scrutinise. Question 17 asked users who did inspect the personalisation, what their main reason was for this. Questions 16 and 17 allowed users to enter an answer in their own words if they selected the *Other* option. Question 18 allowed users to provide additional free-form comments. The pre and post test questions are included in Appendix 8.4.1 and 8.4.2.

The user was allowed to complete the post-test any time after completing their pre-test and was able to access the course following the post-test. However, users could not access the pre or post test once they had completed it.

Table 6.3 – Post-Test Qualitative Survey questions.

	Question	Answers
1-10	Post-test quiz questions	-
11	Recall Henrik's hints, the comments in red font which appeared throughout the course material. Which statement describes how you felt about Henrik's hints?	A. They were annoying. I wanted to get rid of them but did not know how. B. They were annoying, I knew how to get rid of them but didn't feel compelled to do so. C. I noticed them but they were NOT annoying. D. I did not notice Henrik's hints.
12	Recall the jokes that appeared in blue font throughout the course material. Which statement describes how you felt about them?	A. They were annoying. I wanted to get rid of them but did not know how. B. They were annoying, I knew how to get rid of them but didn't feel compelled to do so. C. I noticed them but they were NOT annoying. D. I did not notice the jokes.
13	I knew how to inspect the personalisation to see what items were included or removed from my personalised view.	A. Strongly Agree B. Agree C. Neutral D. Disagree E. Strongly Disagree
14	I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.	*As above.*
15	I feel it is useful to be able to inspect and control the personalisation.	*As above.*
16	If you did NOT inspect the personalisation, what was the main reason.	A. Did not feel compelled to B. I trusted the system's default personalisation C. Other *(allowed free form comments entry)*
17	If you DID inspect the personalisation, what was the main reason.	A. Curiosity B. I did not like how a page was personalised to me C. A belief about me was wrong D. Other *(allowed free form comments entry)*
18	Do you have any other comments?	*Free form comments entry.*

SASY Content

The course covered three modules: the UNIX Shell, UNIX File System, the Chmod command. Each module consisted of several pages that introduced concepts, followed by several pages of quiz questions. The course contained 39 pages in total and required at least one hour to work through.

On logging in the user was presented with the initial profile page, discussed in the next section. Next the user completed the pre-test and navigated to the course Home page. This page adaptively displayed navigation elements based on the user's knowledge, as specified in their user model. If the student did not already know a module, the home page provided hyperlinks to teaching pages for that module.

Otherwise the page provided links to the quiz pages for that module. Once the student had read through the teaching pages and quiz pages for a module, the Home page provided links to the next module. So the home page was adapted with links hidden or displayed, depending on the assessed progress of the student. When the student completed all three modules they were presented with an exercise task that showed them how to create a batch shell script, create a web page or write a batch script from the web, based on their selection on the initial profile page. To change the task that was presented by SASY, the user had to change their profile.

Personalisation added several content elements to pages throughout the course if a user's profile indicated they wanted it. These elements were hints, jokes, advanced UNIX security concepts, additional quiz questions. Figure 6.1 includes a screen shot of a typical content page from the UNIX topic. The top of the page includes a UNIX joke. Also note 'Henrik's Hint' included toward the bottom of the page. These elements were added by the personalisation because the user model indicated the user wanted to see jokes and hints. Notice these user model attributes are displayed on the right hand side of the page as the bullet points *You want Henrik's hints* and *You want UNIX jokes*.

Initial Profile

The initial profile required the user to specify two user model attributes, as documented in Table 6.4. The first column shows the system internal name of the user model attribute, e.g. *Goal*, and the text displayed to the user on the profile page *What task do you want to practice in this tutorial?* The second column lists all the possible answers from which the users had to select one response, for example, *Learn how to create and run a batch script*. The third column describes the effect the response had on the personalisation. The *Goal* attribute controlled which exercise task the user was shown, as described above. The *Level* attribute did not affect personalisation and was used for information only to track our user population. Note that the attribute names, e.g. Goal, were not displayed to users and were only used by the system internally. We use the names in this paper to allow us to refer to the user model attributes easily.

Table 6.4 – User model attributes displayed on initial profile page.

Attribute	Value	Effect on Personalisation
Goal What task do you want to practice in this tutorial?	Learn how to create and run a batch script	Controlled which link was shown on the home page and course map page. To change which link was shown, the user had to change this profile attribute.
	Learn how to make a web home page	
	Learn how to run a script from the web	
Level What course are you in?	First year computer science student	Had no effect on personalisation. Was used purely to track the background of our student population.
	SOFT2004 student	
	UIDP student	
	Third year computer science student	
	Summer School student	
	None of the above	

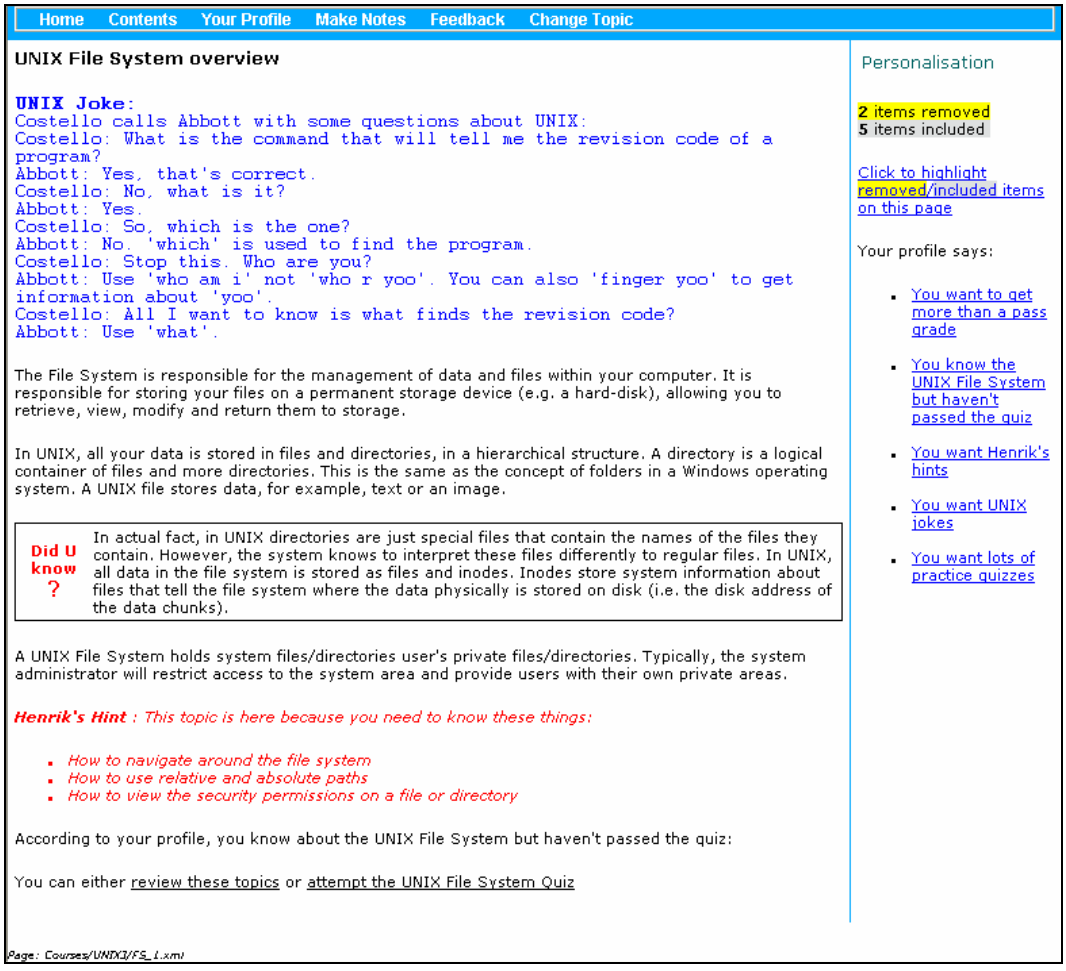


Figure 6.1 – Sample of typical content page in UNIX topic.

[Top of the page includes a UNIX joke. Also note 'Henrik's Hint' included toward the bottom of the page. This content was added because the user model indicated the user wanted to see jokes and hints.]

Complete Profile

If a user returned to the profile page after initial profile creation, the complete profile page was displayed: it included user model attributes that had been observed or inferred about the user. Inferred attributes are described in Table 6.5. A screen shot of the complete profile is shown in Figure 6.2 (page 154). The first column shows the system internal name of the user model attribute, e.g. *Jokes*, and the text displayed to the user on the profile page *Do you want the odd UNIX joke included in the notes?* The second column lists all possible selections, for example, *Yes* or *No*. We have also indicated the SASY default in this column. For example, by default SASY assumed users wanted to see jokes in their course material. The third column describes the effect the response had on the personalisation. For example, selecting *Yes* for the *Jokes* attribute caused jokes to be added on all pages in the course.

There were three attributes that inferred the user's knowledge, corresponding to the three modules: *knowShell*, *knowFileSys* and *knowChmod*. Each of these was set by completing the pre-test as follows. If the student correctly answered all questions relevant to a module in the pre-test, SASY set the corresponding profile attribute to *You probably know this but have not passed the quiz on it*. This meant the student would not have to read through the teaching material for the module, but would still have to complete a quiz on the topic to obtain full credit for it. However, if the student got one question wrong in the pre-test, SASY set the profile attribute to *No* claiming the student had no knowledge about the module at all. This meant the student would have to access all the pages of the teaching material and complete the quiz for the module, even though they probably already knew about it.

Table 6.5 – Inferred user model attributes.

Attribute	Value	Effect on Personalisation
Learning Preferences		
Jokes Do you want the odd UNIX joke included in the notes?	Yes – show a few jokes (DEFAULT)	Jokes were added on all pages in the course.
	No	No jokes were added on pages.
Hints Do you want to see Henrik's Hints?	Yes – show hints (DEFAULT)	Additional comments about concepts were added on all pages in the course.
	No – DO NOT show hints	No hints were added on pages.
Grade What level of knowledge do you want to achieve through this course?	Learn just enough to pass	Only simple concepts were presented and quizzed.
	Learn as much as you can to get more than a pass grade (DEFAULT)	Additional advanced pages and concepts were presented. Quiz questions were more difficult.
Assessed How do you want to be	Lots of practice quizzes (DEFAULT)	The student was shown 6 pages of quiz questions per module and had to correctly answer 6 questions correctly to pass a module.

Attribute	Value	Effect on Personalisation
assessed?	A few practice quizzes	**Two pages** of quiz questions shown per module, **2** questions answered correctly to pass a module.
	No practice quizzes	The student was not shown any quiz questions. Once they 'read' a module SASY assumed they knew it.
Knowledge		
KnowShell Do you know about the UNIX Shell?	No – never heard of it (DEFAULT if user answered even one Shell question incorrectly in pre-test)	The course home page directed the student to access pages to learn about UNIX Shell.
	You probably know this but have not passed the quiz on it (DEFAULT if user answered ALL Shell questions correctly in pre-test)	The course home page directed the student to access pages to complete the UNIX Shell quiz, provided they wanted to be quizzed.
	Yes – you have passed the quiz on it	The course home page directed the student to access the page to learn about their selected task, provided they had passed the quiz on the other topics.
KnowShellQuiz The total number of questions you answered correctly about the UNIX Shell	Current Quiz Score (DEFAULT 0)	Used in conjunction with the Assessed attribute to determine whether a user had passed the UNIX Shell Quiz. If not, SASY directed users to the page to learn about the Shell and repeat the quiz.
KnowChmod KnowChmodQuiz	Same as above two rows but in relation of knowledge about UNIX Chmod command.	
KnowFileSys KnowFileSysQuiz	Same as above two rows but in relation of knowledge about UNIX File System.	
System Preferences		
Show By default, do you want SASY to highlight the personalised items on each page?	Yes	On each page, the content that was added and removed by the personalisation was highlighted by default. That is, the Highlight tool was always invoked by default
	No (DEFAULT)	Content was not highlighted by default.
ShowBeliefs By default, do you want SASY to show you beliefs about you that affected personalisation on each page?	Yes (DEFAULT)	The list of profile attributes that affected personalisation on the page were displayed in the right hand panel of the page.
	No	Profile attributes were not displayed on the page – making the SASY page like a conventional adaptive hypertext.

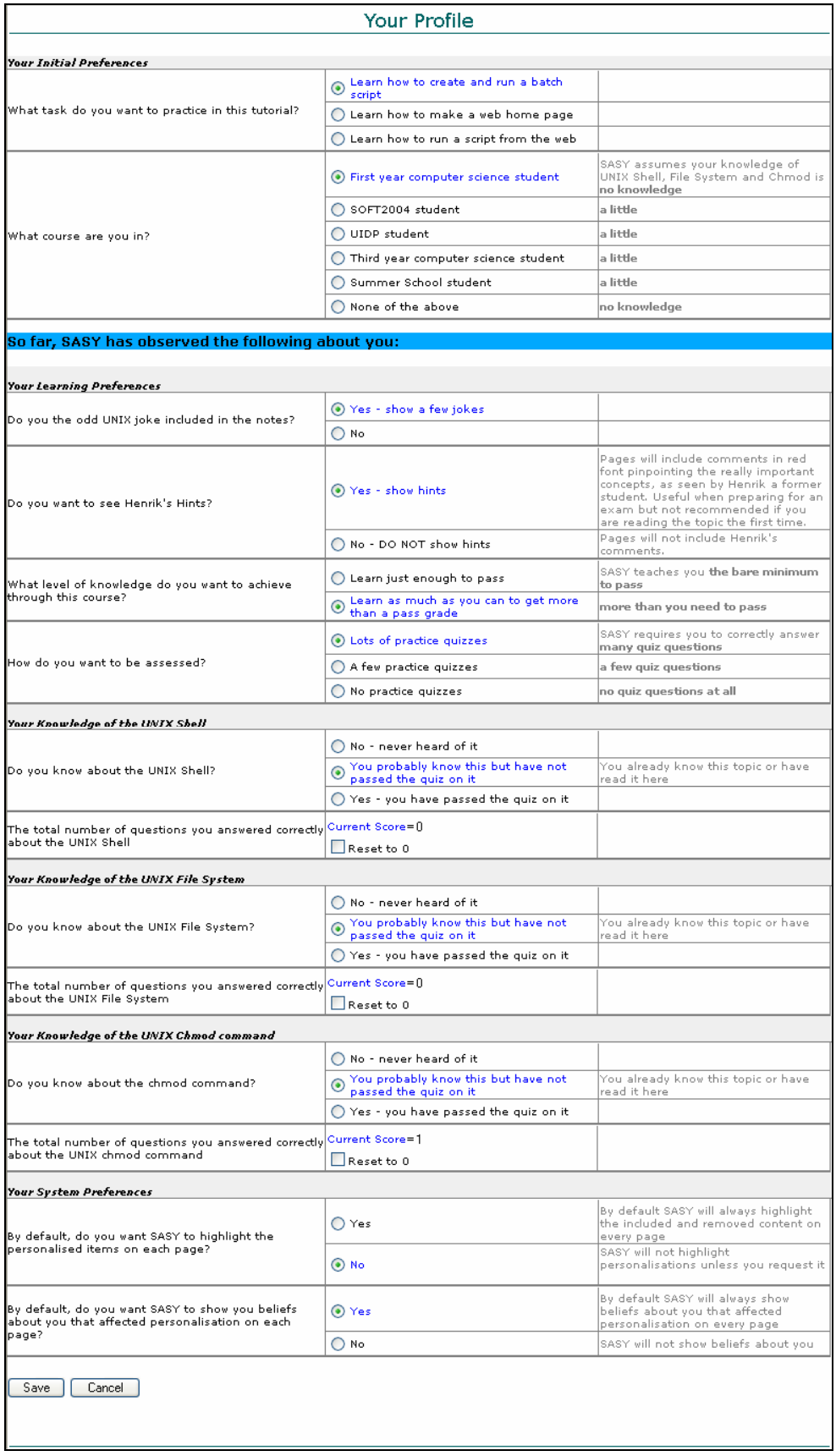


Figure 6.2 – Screenshot of the profile page showing the complete profile with initial preferences and inferred attributes.

Other SASY Topics

In addition to accessing the UNIX Security course, users were free to select other SASY topics, but were not asked to do so. The other SASY topics were:

- Guide to Personalisation in SASY – a two page guide on using the SASY scrutinisation tools.
- Holiday Planner – the content for a different evaluation that presented a personalised holiday recommendation system.
- TV Guide – the content for a different evaluation that presented a personalised television guide.

Usage of these SASY topics is discussed in the results section.

6.2.4 Results

We firstly report overall use of SASY scrutinisation tools in terms of total usage time, total invocations of each tool, types of scrutinisation performed by users and their timings. Next we analyse the usage of each of the scrutinisation tools: Evidence tool, Highlight tool, Profile tool. Finally we present the results of the qualitative survey and additional feedback.

The Appendix includes additional graphs that illustrate the usage of the tools, and a table that summarises the scrutinisation of each user. We refer to the appendix as appropriate. The Appendix also contains an analysis that could not be presented here for the lack of space, describing the scrutinisation of the top five users of scrutinisation tools.

Total Usage Time

First we analysed the total time users spent using SASY and the number of sessions they had. Since users were not under direct observation, we have no way of knowing whether a user was actively reading a page at any given time or whether they had walked away from their terminal. Therefore, we assumed that 60 minutes of inactivity indicated the end of a session. It is reasonable to expect a few minutes of inactivity between user actions as users may have been practicing concepts learnt in a UNIX terminal whilst working through the course. A summary of usage time and the number of sessions is shown in Table 6.6. On average, students spent about one hour using the system during two sessions. Overall, this indicates serious interaction time with SASY. At the extremes, there were 6 users who spent less than 10 minutes and one user spent 3.5 hours using SASY over 4 sessions. Studying the log file for the user who spent 3.5 hours using SASY (user 103, see appendix 8.4.3, page 240), the logged actions show a recurring behaviour of reading several pages, pausing for up to ten minutes, then continuing. Perhaps this user was using SASY whilst simultaneously doing something else, perhaps trying out the UNIX commands. This user completed the pre and post tests in the second session, but returned to SASY one day and then four days later to work through the content and quiz questions again.

A graph of the total usage time, with users on the x-axis ranked least usage to most is shown in Appendix 8.4.5 (page 244). Appendix 8.4.6 (page 245) shows a histogram of the frequency of session start hour. It shows the bulk of sessions started during morning hours (9am-12pm) and evening hours (9pm-12am).

Table 6.6 – Summary of Usage, Pre and Post test scores.

	Total Usage Time (HH:MM:SS)	No. Sessions	Pre-Test Scores	Post-Test Scores
Minimum	0:02:34	1	1	5
Maximum	3:32:49	5	10	10
Average	1:01:05	2	6.31	8.83
Median	0:50:48	2	6	9
Standard Deviation	0:42:23	1	2.34	1.38
Sample (N)	N=105	N=105	N=105	N=82

Learning Gains

Although this experiment did not evaluate learning gains due to usage of SASY, on average, post test scores (avg. 8.83) were higher than pre-test scores (avg. 6.31). The statistics are shown in Table 6.6. Users were not supervised so results may not be highly reliable.

We did not find any correlation between the learning gain and use of the scrutinisation tools. We did receive anecdotal evidence, that SASY was effective in that university class tutors commented that students did far better than usual in the lab class related to the SASY homework assignment, and generally had fewer UNIX access control problems than in previous years. This might be due to the SASY teaching material, and not necessarily the scrutinisation tools.

Scrutinisation Usage Summary

We examined the number of scrutinisation actions performed by users. Table 6.7 summarises the usage of the scrutinisation tools. The table summarises the total number of scrutinisations made by users (A) and provides a breakdown for each method of invoking the Profile tool (B,C) and Profile tool combined total (D), Evidence tool (E,F, combined total G) and Highlight tool (H). It also includes the number of user model attributes that were changed by users in their profile (I). The statistics in Table 6.7 include only scrutinisation actions performed in the UNIX course. Scrutinisation of the other SASY topics is discussed later. In addition, the usage for each user is included in Appendix 8.4.3 (page 240).

Table 6.7 – Summary of Scrutinisation Tool usage

Key:

(A) The total number of scrutinisations by a single user. For example, one user made 22 scrutinisations, 81 (77%) users scrutinised at least once.

(B) Number of times users opened the Profile tool from the menu at the top of a page.

(C) Number of times users opened the Profile tool from the Evidence tool.

(D) Number of times users opened the Profile (total of B and C).

(E) Number of times users accessed the Evidence tool from a content page.

(F) Number of times users accessed the Evidence tool from the profile page.

(G) Number of times users accessed the Evidence tool (total of E and F).

(H) Number of times users accessed the Highlight tool.

(I) Number of profile attributes changed by a user in their profile.

Row	N=105	(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(D) Opened Profile (total)	(E) Viewed Evidence from Regular Page	(F) Viewed Evidence from Profile Page	(G) Viewed Evidence (total)	(H) Used Highlight Tool	(I) Changed Profile attributes
	Statistics									
1	**Min.**	0	0	0	0	0	0	0	0	0
2	**Max.**	22	5	3	5	7	4	7	8	11
3	**Average**	3.69	0.76	0.15	0.91	0.78	0.16	0.94	0.67	1.16
4	**Median**	3	0	0	0	0	0	0	0	0
5	**Std. Dev.**	4.14	1.11	0.5	1.29	1.29	0.62	1.50	1.3	2.07
	No. Users									
6	**>= 1 use**	81 (77%)	46 (44%)	11 (10%)	49 (47%)	41 (39%)	8 (8%)	45 (43%)	36 (34%)	37 (35%)
7	**> 2 uses**	55 (52%)	9 (9%)	1 (1%)	11 (10%)	8 (8%)	2 (2%)	11 (10%)	8 (8%)	18 (17%)
8	**> 4 uses**	23 (22%)	2 (2%)	0 (0%)	4 (4%)	3 (3%)	0 (0%)	6 (6%)	3 (3%)	7 (7%)
9	**> 8 uses**	13 (12%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)

From a statistical view of overall scrutinisation tool usage, the results indicate substantial use of the scrutinisation tools. Column A of Table 6.7 shows 81 users (77%) scrutinised in some way, 55 (52%) scrutinised more than twice, 23 (22%) scrutinised 5 or more times (i.e. more than average), 13 (12%) scrutinised 9 or more times (i.e. more than twice average). However, the usage of each tool separately is not so high (columns B to H). Each tool was used by less than half of the users and the median number of times a tool was used is zero. The result was unsurprising, given the user's main task was to learn about UNIX security. Any scrutinisation was additional distraction from their core task. Indeed, a significant part of the motivation for a field trial was to enable us to observe the proportions of users who chose to not scrutinise and the proportion who only use some of the tools.

The profile was mainly accessed from the menu at the top of a page rather than from the evidence tool. The evidence tool was mainly accessed from a regular page rather than from the profile. Some users accessed the profile and evidence tool both ways. The highlight tool could only be accessed from regular pages and is not accessible from the profile.

Column I shows that there was a real interest among users in exercising control over the personalisation by changing their profile. This is evident in the results in Column I, rows 7, 8 and 9.

Next we examined the statistics with non-users removed. That is, for each tool we removed the actions of users who did not use the tool. The statistics are reported in Table 6.8. This enables us to see the average behaviour of those who did scrutinise. Row 1 shows the sample size, that is, the number of users who used the tool (same as row 6 in Table 6.7). Rows 2 and 3 show the average and median values, with the difference from the corresponding value in Table 6.7 shown in parenthesis. This indicates that users who did use a scrutinisation tool at all, on average used it more than once.

Table 6.8 – Summary of Scrutinisation Tool usage with non-users removed. For column key refer to Table 6.7.

Row		(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(E) Viewed Evidence from Regular Page	(F) Viewed Evidence from Profile Page	(H) Used Highlight Tool	(I) Changed Profile attributes
1	**Sample Size (N)**	81	46	11	41	8	36	37
2	**Average**	4.78 (+1.09)	1.74 (+0.98)	1.45 (+1.3)	2.0 (+1.22)	2.13 (+1.97)	1.94 (+1.27)	3.30 (+2.14)
3	**Median**	4 (+1)	1 (+1)	1 (+1)	2 (+2)	2 (+2)	1 (+1)	2 (+2)

Graph of total scrutinisation tool usage

We graphed the total number of scrutinisation actions users performed in Figure 6.3, with users ranked from least usage to most number of scrutinisations. From this graph it is clear a small number of users used the tools extensively. This is consistent with Table 6.7 (Row 9) that showed 13 (12%) scrutinised 9 or more times, more than twice the average.

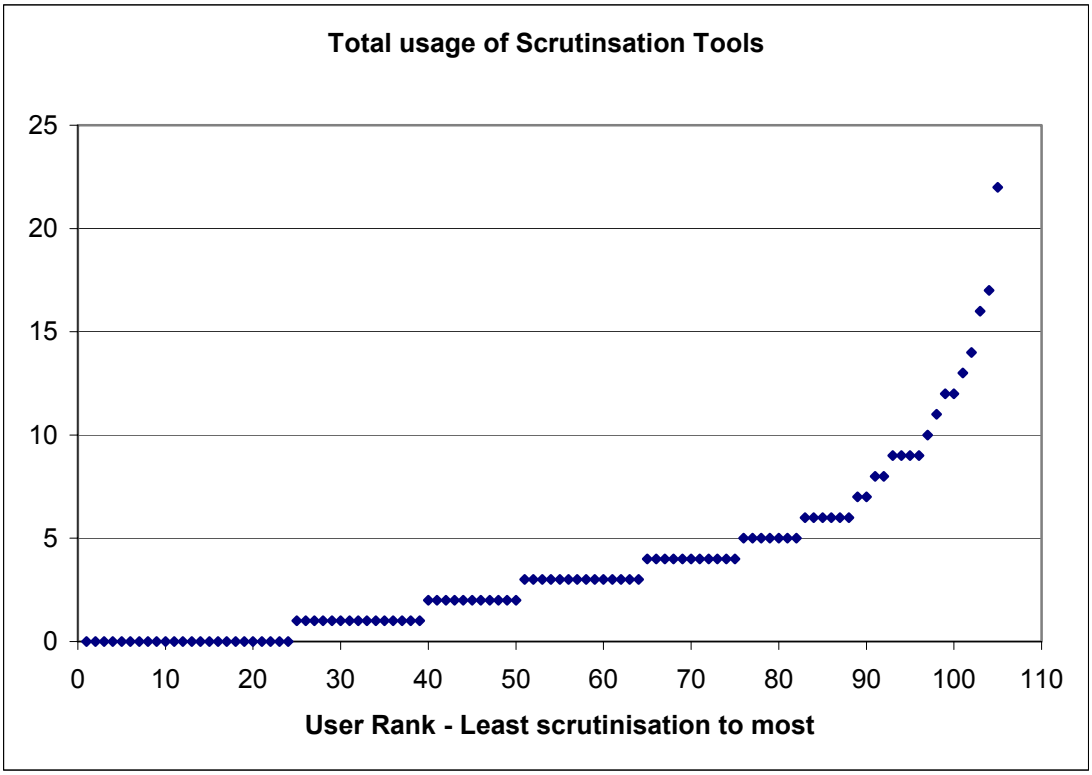


Figure 6.3 – Graph of total usage of scrutinisation tools, with users ranked least use to most.

Forms of scrutinisation

We constructed a histogram (Figure 6.4) to get the frequency of the four forms of scrutinisation: use of Profile tool, use of Evidence tool, use of Highlight tool, changing profile. That is, for each student we counted how many forms of scrutinisation they used. The resulting left-skewed distribution shows the largest single group of students (26.67%) used two forms of scrutinisation (average = 1.59, median = 2, std. dev. = 1.20). There was a strong correlation (coefficient=0.76) between the number of forms of scrutinisation and the total scrutiny actions for users, meaning that users who scrutinised more overall tended to use more forms of scrutinisation.

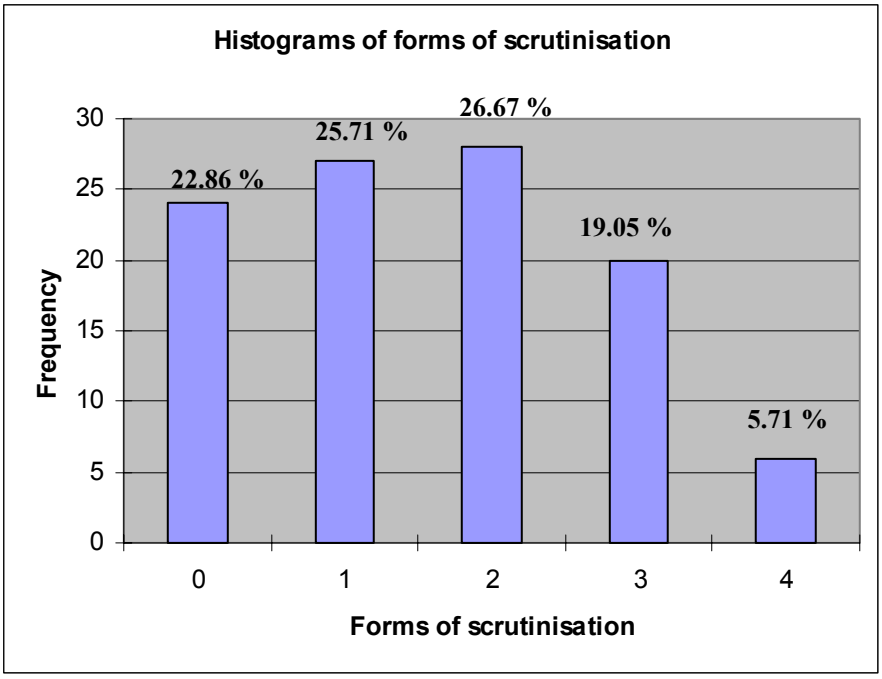


Figure 6.4 – Histogram showing how many types of scrutiny tools were used.

Correlation between forms of scrutinisation

We also calculated correlation coefficients between each form of scrutinisation, but found there was no strong correlation (> 0.5) between the different forms of scrutinisation in Table 6.7.

However, the co-efficient of correlation between usage of the profile tool and number of access to the course map was significant (0.63). Usage of the course map was substantial: 38 (36%) users accessed it at least once (min = 0, max = 5, average = 0.73, std. dev. = 1.23). The course map is a different kind of scrutinisation tool, based on an adaptive navigation guidance technique used by ELM-ART (Weber, Brusilovsky, Schwarz, 1996; Weber, Brusilovsky, 2001). It uses colour to annotate hyperlinks to show the user how they have progressed through the course and it changes as the user visits pages. Since it is not the focus of this thesis, it is not included in Table 6.7.

Timing of scrutinisation

Next we analysed when users scrutinised. We noticed from the logs that users had scrutinised and used the course map at similar points in their session. To express this formally, we assigned each scrutinisation action to one of the following timing categories, to capture the point at which the scrutinisation occurred during a session:

- Start of Session – towards the start of a session after logging in or before accessing the pre-test.
- After Pre-test – immediately after completing the pre-test.

- After Post-test – immediately after completing the pre-test.
- After module – immediately after reading the teaching material or completing the quiz for a module (File System, Shell, Chmod command).
- End of Session – towards the end of the users session in SASY.
- Other – all other times such as in the middle of reading module teaching material or a module quiz page.

The timing results for scrutinisation actions (from Table 6.7) are shown in Table 6.9. The table also shows accesses to the course map (column J) and the number of users (column K). For example, the first row shows 108 (27.91%) scrutinisation actions occurred just after the completion of a learning module. Of these actions, 24 were accesses to the profile from the menu, and so on. In addition, 14 actions were accesses to the course map from the menu. In total, these 108 actions were performed by 32 users.

The table shows most scrutinisation actions occurred when the students were either about to start, or had just finished a learning module, session or pre/post test. Only 14 actions (3.62%) fall in to the Other category so not many students scrutinised while they were learning. This also applies to the use of the course map. In fact, this result also shows the course map was used around the same time that users scrutinised.

Table 6.9 – Timing of Scrutinisation actions. For column key refer to Table 6.7.

Timing	(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(E) Viewed Evidence from Regular Page	(F) Viewed Evidence from Profile Page	(H) Used Highlight Tool	(I) Changed Profile attributes	(J) Accessed Course Map	(K) No. Users
After Module	108	24	2	24	8	16	34	14	32
Start of Session	90	12	3	10	5	16	44	10	38
After Post-test	89	22	7	23	0	15	22	17	27
After Pre-test	45	6	2	12	0	9	16	0	22
End of Session	41	13	2	12	4	7	3	11	23
Other	14	3	0	1	0	7	3	1	8

We summarise our observations so far of scrutinisation tool usage:

Observation 1: Overall, use of the scrutinisation tools was quite high given the context they were using SASY to learn about UNIX Security and any scrutinisation was an additional to their core learning activity: 81 users (77%) scrutinised at least once, 55 (52%) scrutinised more than twice, 13 (12%) scrutinised 9 or more times (i.e. more than twice average).

Observation 2: Different users scrutinised in different ways. For example, some users only used the Profile tool, some only used the Highlight tool, while some used all the tools. This observation comes from the fact there was no correlation between forms of scrutinisation explored by users. It is also evident in the table included in Appendix 8.4.3 (page 240) that shows the number of times each user used each scrutinisation tool.

Observation 3: Overall, users used more than one form of scrutinisation. The largest single group of students (26.67%) used two forms of scrutinisation (average = 1.59, median = 2, std. dev. = 1.20).

Observation 4: There was a strong correlation (coefficient=0.76) between the number of forms of scrutinisation and the total scrutiny actions for users, meaning that users who scrutinised more overall tended to use more forms of scrutinisation.

Observation 5: Overall, users who used a scrutinisation tool at all, used it more than once, based on the results in Table 6.8.

Observation 6: There was a strong correlation (co-efficient =0.63) between usage of the profile tool and course map.

Observation 7: Most usage of the scrutinisation tools and course map occurred when the students were either about to start, or had just finished a learning module, session or pre/post test, based on timing results in Table 6.9.

We also analysed in detail, the scrutinisation actions of the five users that scrutinised most. The analysis is reported in Appendix 8.4.4, page 242. It shows some of the quite different ways and times that users scrutinised.

Next we examine usage of the Evidence tool, Highlight tool and Profile tool in detail.

Usage of Evidence Tool

We analysed the logs to determine which profile attributes appeared to trigger exploration of the underlying evidence. The results are shown in Table 6.10. Users mostly sought evidence about knowledge attributes (UNIX File System had 21 users, Chmod command had 16 users, UNIX Shell had 9). There were also 3 users who sought evidence for their current score in UNIX Shell module quiz. In all, 37 users sought evidence about at least one of these three knowledge attributes, 6 of these sought evidence for more than one. This indicates that knowledge was the most popular user model attribute type for which evidence was sought. Perhaps users wanted to know why SASY thought they did or did not know something, or perhaps wanted to know how they were doing – a skillometer might have shown that.

Users also sought evidence for non-knowledge attributes. The Assessed attribute controlled how many quiz questions SASY would show the user. There were 15 users who sought evidence of this attribute. Five users also sought evidence for the Grade attribute, that caused the personalisation to present students with advanced concepts. Only 2 users sought evidence as to why SASY thought they wanted jokes in the teaching material and no users sought evidence about the Hints attribute. Nine users sought evidence for the Goal attribute, that was specified by users in the initial profile.

Overall, users were more concerned with why SASY stated they did or did not know something than why SASY thought they wanted to see jokes, hints, advanced concepts and many quiz questions.

A graph showing the usage of the evidence tool, with users ranked least usage to most is included in Appendix 8.4.7 (page 246). Like the graph for overall scrutiny, some users made little use of the tool while a smaller number of users made heavy use of the tool.

Table 6.10 – Number of users who sought evidence by profile attribute.

Profile Attribute	No. Users who sought evidence about it
Hints – controlled whether students were shown hints	0
Current score for UNIX File system Quiz	1
Jokes – controlled whether students were shown jokes	2
Current score for UNIX Shell Quiz	3
Grade attribute – controlled whether students were shown advanced concepts	5
Goal attribute – set in the initial profile, controlled the exercise task presented	9
Knowledge about UNIX Shell	9
Assessed attribute – controlled how many quiz questions were shown.	15
Knowledge about UNIX Chmod command	16
Knowledge about UNIX File System	21

Timing of Evidence Tool Usage

Next we analysed the timing of the evidence tool usage to determine when users sought evidence for different attribute types. The results are shown in Table 6.11. For all attributes, evidence was mostly sought after a learning module or after the post test. For knowledge attributes, evidence tool usage was also high just after the pre-test. This was expected because SASY changed knowledge attributes after a user completed the pre-test or a learning module and successfully demonstrated some knowledge. The behaviour where users sought evidence after the pre/post test of a learning module is possibly evidence of reflection about their learning. Also, during the main learning activity they probably need to concentrate on the demanding task of learning. They may have only felt able to deflect their activity to scrutinisation at these starting and stopping points.

For non-knowledge attributes (jokes, grade, assessed), however, users did not scrutinise at the start of the session. Perhaps this is because students did not experience the impact of the grade and assessed user model attributes on personalisation until they were presented with a quiz after reading a module, when they were presented with many quiz questions on advanced concepts.

Table 6.11 – Timing of Evidence tool usage. * Knowledge attributes include UNIX File System, Chmod command and Shell.

Timing	Number of times evidence tool was invoked				
	Knowledge* Attributes	Jokes Attribute	Grade Attribute	Goal Attribute	Assessed Attribute
After Module	17	0	5	5	5
Start of Session	12	0	1	0	2
After Post-test	9	0	2	7	5
After Pre-test	11	0	0	0	1
End of Session	10	2	0	2	2
Other	1	0	0	0	0

Pages where Evidence Tool was used

Out of the 99 times the evidence tool was invoked, it was mostly invoked from the home page (69 times, 69.70%), followed by the profile page (17 times, 17.17%). It was also invoked twice from the course map page, once from the last page of the File System module, once from the last page of the Chmod module. This might be due to the cognitive load of their core learning activity.

Other observations on the use of the Evidence Tool

As mentioned in the design of this experiment, we had expected some users might scrutinise why SASY claimed they did not have any knowledge about a module thought they demonstrated some knowledge in the pre-test, but did answer every question correctly. As shown in Observation 8, this appears to have been the case.

Observation 8: Some users sought evidence about their alleged lack of knowledge about a concept immediately after completing the pre-test, even though they demonstrated some level knowledge in the pre-test by correctly answering some, but not all, questions about that concept [11 (10%) users (40, 42, 47, 51, 56, 69, 76, 82, 85, 94, 105)].

Observation 9: Some users sought evidence of a knowledge attribute immediately after manually increasing the knowledge level in profile. These users seemed to want to test whether SASY had actually let them change their profile without contest [3 (3%) users (92, 95, 98)].

Observation 10: Some users sought evidence of an attribute immediately after logging in for the first time, before attempting the pre-test. These users noticed the personalisation column on the SASY page and probably wanted to know how SASY had inferred information about them [3 (3%) users (53, 73, 83)].

Usage of Highlight Tool

Timing of Highlight Tool Usage

Similarly to the usage of the evidence tool, the highlight tool was most often invoked either just before/after completing the pre-test, after completing the post-test or learning module, as shown in Table 6.9 (page 161).

Pages where Highlight Tool was used

Out of the 70 times the Highlight tool was invoked, it was mostly invoked from the home page (55 times, 78.57%), followed by the course map page (4 times, 5.71%).

Other observations on the use of the Highlight Tool

Observation 11: Some users used the Highlight tool on home page and attempted to repeat pre-test or post-test [7 (7%) users (39, 58, 59, 68, 84, 87, 103)].

Once the user completed a pre-test or post-test, the hyperlink to the test was subsequently removed by the personalisation. We noticed that 7 users used the Highlight tool and accessed the links that had been removed by the personalisation. These 7 users first invoked the highlight tool to display the removed links, then clicked the link to attempt to repeat the pre-test or post-test. There was some feedback provided by students, outside of the evaluation, that they did not realise they would not be able to retrieve their post-test results in SASY at a later time. A print out of the post test was required to be handed in at their class as proof they had completed their homework. It is possible some of these seven users were attempting to repeat the post test to prove they had completed their homework. This also showed they knew the Highlight tool could be used to see things that had changed or had been removed by personalisation.

Observation 12: Some users changed a system preference in their profile so that the highlight tool was invoked on every page by default [7 (7%) users (82, 89, 93, 95, 96, 102, 105)].

Interestingly, 7 (7%) users changed the system preference in their profile to highlight personalisation on every page by default. Of these, 2 users changed it back to the original setting soon after trying it

out and 3 users worked in this mode for the whole course or a substantial part of it. This is surprising as we assumed it would be distracting to have personalisation highlighted on every page by default. One user who exhibited such behaviour later explained why they did this: "I asked it to show personalisation since I wanted to make it more readable and to highlight the things of interest to me. I was also simply interested in how the system worked, so I wanted to see how the system adapted to my needs and wants".

User 102 exhibited particularly interesting behaviour. At the end of their session having completing the post-test, they accessed their profile from the Home page and reset their quiz scores and changed the system preference attribute to highlight personalisation on every page by default. The user then worked in this mode for the rest of their session (about an hour), reading through the chmod quiz section. It seems the user wanted to ensure they saw all the available quiz questions and therefore needed to view the non-personalised view of each page.

A graph showing the usage of the Highlight tool, with users ranked least usage to most is included in Appendix 8.4.8 (page 247).

Usage of Profile Tool

Table 6.12 shows profile attributes that were changed by users. The second column shows the number of users who changed each attribute. The third column shows the number of users who also sought evidence for that attribute using the Evidence tool. The table shows that most of the users who made changes to their profile did not use the evidence tool to first investigate why and how the profile attribute was set by SASY before changing it. Hence there is little correlation between the users listed in Table 6.10 and Table 6.12.

Table 6.12 – Number of users who changed profile attributes.

Profile Attribute	No. Users who changed it	No. Users who also sought evidence about it
Grade attribute – controlled whether students were shown advanced concepts	1	0
Reset score for UNIX Chmod Quiz	2	0
Reset score for UNIX File System Quiz	3	0
Reset score for UNIX Shell Quiz	4	0
Hints – controlled whether students were shown hints	4	0
Assessed attribute – controlled how many quiz questions were shown.	6	2

Profile Attribute	No. Users who changed it	No. Users who also sought evidence about it
System preference to Highlight personalisation on each page by default	7	0
Jokes – controlled whether students were shown jokes	8	0
Goal attribute – set in the initial profile, controlled which exercise task was presented	15	4
Knowledge about UNIX File System	15	1
Knowledge about UNIX Shell	15	0
Knowledge about UNIX Chmod command	17	4

There were 4 users (4%) who changed their profile to remove Hints, 8 (8%) removed Jokes, 1 user changed the Grade attribute to remove advanced concepts and 6 (6%) changed their profile to reduce the number of quiz questions. One user initially decreased the number of questions then increased it to the original setting. Also 15 users changed the Goal attribute to have SASY present a different task, as described in Table 6.4 (page 151).

The most commonly changed profile attributes related to knowledge about the UNIX File System, UNIX Shell and Chmod Command: 13 users changed their profile to state they knew *more* about the File System and the Shell than indicated by SASY; 2 users stated they knew less; 15 users stated they already knew about the Chmod Command, 2 users stated they knew less.

Timing of Profile Tool Usage

Like the usage of the Evidence tool and Highlight tools, the profile was most often invoked either just before/after completing the pre-test, after completing the post-test or learning module, as shown in Table 6.13. The table shows that unlike the results for the evidence tool (Table 6.11), profile attributes were often changed at the beginning of a session or just after the pre-test. This supports the observations mentioned above. This is true for both knowledge and non-knowledge attributes (jokes, grade, assessed). It seems users were comfortable to take control of the personalisation by changing their profile at the start of their session to remove jokes, hints and reduce the number of quiz questions. When users logged in for the first time they were only presented with the initial profile. However, if they accessed the profile again by logging in again, or explicitly invoking the profile tool the complete profile page was displayed showing initial preferences and all inferred attributes. It is possible that simply seeing the inferred attributes (that were previously hidden) might have encouraged some users to change them.

Table 6.13 – Timing of profile attribute changes. * Knowledge attributes include UNIX File System, Chmod command and Shell.

Timing	Number of times user model attribute was changed					
	Knowledge* Attributes	Jokes Attribute	Hints Attribute	Grade Attribute	Goal Attribute	Assessed Attribute
After Module	14	4	2	0	6	5
Start of Session	26	2	2	0	8	0
After Post-test	15	0	0	1	3	2
After Pre-test	10	2	0	0	1	1
End of Session	0	0	0	0	3	0
Other	0	0	0	0	2	0

We also made the following timing-related observations.

Observation 13: Some users changed their profile to state they knew *more* about a module that what SASY said they did immediately after logging in, before attempting the pre-test [5 (5%) users (75, 67, 68, 77, 101)].

Observation 14: Some users changed their profile to state they knew *more* about a module than what SASY said they did, just after completing the pre-test. They demonstrated some level of knowledge in the pre-test by correctly answering some, but not all, questions about that concept [5 (5%) users (66, 72, 82, 93, 105)].

Observation 15: Some users changed their profile to state they had *less* knowledge than indicated by SASY, having completed the whole course once already. These users seemed to be revising the course or wanted to repeat the teaching experience [2 (2%) users (52, 94)].

Observations about whether users scrutinised to predict the impact of profile changes

Observation 16: Overall users did not use the Evidence tool before making profile changes. However, some users did so [4 (4%) users (79, 81, 94, 105)].

Observation 17: Overall users did not use the Highlight tool to investigate the effect of profile changes before making changes. However, one users did so [1 (1%) users (81)].

A graph showing the usage of the Profile tool, with users ranked least usage to most is included in Appendix 8.4.9 (page 248).

Usage of Other SASY topics

As mentioned previously, students who accessed the UNIX security course were free to access other SASY topics but were not asked to do so. There were 39 users who accessed the SASY Introduction Guide, 12 accessed the Personalised TV Guide, 12 accessed the Holiday planner.

The SASY introduction guide described the SASY scrutinisation tools and asked users to practice using the tools. Out of the 39 users who accessed the guide, 30 (77%) scrutinised in the UNIX course. However, 51 users who did not access the SASY Introduction guide did scrutinise in the UNIX course. The correlation co-efficient for the number of scrutinisation actions users performed in the UNIX course against those performed in the SASY Introduction guide was 0.29, indicating there was very little correlation.

Table 6.14 shows the number of users who scrutinised in the SASY Introduction Guide, TV Guide and Holiday Planner topics. As expected, there was high use of the scrutinisation tools in the SASY introduction guide: it directed the user to do so. However, users also scrutinised in the other topics out of their own free will. This could be a reflection of the curiosity of users about the personalisation. There was no correlation between the number of times users scrutinised in the UNIX course versus that of other SASY topics. The order in which topics were listed when the user logged in was SASY Introduction Guide, UNIX Security, Holiday planner and lastly the Personalised TV Guide, which may have had some impact in terms of which topics were accessed.

Table 6.14 – Number of users who scrutinised in the SASY Introduction Guide, TV Guide and Holiday Planner topics. For column key refer to Table 6.7. Recall that there were 81 users who scrutinised in the UNIX course. All the users listed in this table had also scrutinised in the UNIX Course.

Timing	(A) Scrutinised in some way	(B) Opened Profile from menu	I Opened Profile from evidence	(E) Viewed Evidence from Regular Page	(F) Viewed Evidence from Profile Page	(H) Used Highlight Tool	(I) Changed Profile attributes
SASY Introduction	14	5	3	4	0	12	5
TV Guide	2	1	0	0	1	1	1
Holiday Planner	5	2	0	1	0	3	1

User Survey

As part of completing the post-test, users were asked five questions about qualitative aspects of their experience using the scrutinisation tools. Although 105 students used the UNIX Security course, 82 completed the post-test, 8 provided free form comments. Of those 82, 10 were excluded from the results: 2 answered A to all qualitative questions which does not logically make sense, one user's results were lost through a technical error, 8 had completed the post-test immediately after the pre-test without using SASY, therefore it would not be fair to include them in the results. This leaves 72 (69%) user responses included in the qualitative survey.

The results for qualitative Questions 11 and 12 (from Table) are shown in Table 6.15. These questions asked users to indicate how they felt about the Hints and Jokes included in the teaching material. Surprisingly, most users did not find the jokes or hints annoying at all or did not even notice them. This explains why only 4 users actually changed their profile to remove Hints, 8 removed Jokes. On the positive side, 5 users said they knew how to remove the hints, 13 knew how to remove the jokes but did not feel compelled to do so.

There were 2 users who said they did not know how to remove the hints and 4 who did not know how to remove the jokes. These answers might indicate some users did not know how to scrutinise the system. However, these responses were not consistent: since the method for removing hints and jokes was the same so the survey responses should have been the same. One user said they did not know how to remove jokes or hints, which is consistent. However, another user said they did not know how to remove the jokes but did know how to remove hints. This is contradictory since the method to remove

the content is the same in both cases. The other 3 users who said they did not know how to remove hints claimed they did not even notice the jokes and 2 of these did actually remove jokes and/or hints prior to completing the survey. Essentially, due to the contradictory survey responses, it is not possible to accurately determine how many users did not know how to change personalisation to remove jokes and hints from content using the results of this survey alone. However, based on the survey results, it seems fair to claim most users did not find the jokes and hints annoying, a small number of users found them annoying and were able to correct the personalisation while a small number of users may not have known how to change the personalisation.

Table 6.15 – Summary of results of survey questions about Jokes and Hints in course material.

N=72	Which statement describes how you felt about the Hints in SASY? (%)	Which statement describes how you felt about the Jokes in SASY? (%)
They were annoying. I wanted to get rid of them but didn't know how.	2 (3 %)	4 (6 %)
They were annoying, I knew how to get rid of them but didn't feel compelled to do so.	3 (4 %)	10 (14 %)
I noticed them but they were NOT annoying.	51 (71 %)	43 (60 %)
I did not notice them	16 (22 %)	15 (21 %)

Questions 13 through 15 (from Table 6.3) on the survey asked students to indicate whether they agreed with the statements:

- I knew how to inspect the personalisation to see what items were included or removed from my personalised view.
- I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.
- I feel it is useful to be able to inspect and control the personalisation.

The results are shown in a histogram in Figure 6.5. For Questions 13 and 14, 37 (51%) strongly agreed or agreed they knew how to inspect the personalisation to see what items were included or removed from their personalised view and that they could change the beliefs on the profile page to change what was included and removed by the personalisation. Of these, 13 had actually used the Highlight tool, 14 had changed attributes in their profile. On the other hand, 11 (15%) disagreed or strongly disagreed. Of these, 7 had used the Profile tool to change attributes, 3 had used the Highlight tool.

For Question 15 in the survey, 43 (60%) strongly agreed or agreed it was useful to be able to inspect and control the personalisation. On the other hand, 7 (10%) disagreed or strongly disagreed. Of these, 5 users had also disagreed or strongly disagreed to one or both of the previous questions. Basically, these users did not see the scrutinisation tools as useful.

The next question asked of users who had *not* inspected the personalisation, what was their main reason: 36% selected *Did not feel compelled to do so*, 31% selected *Trusted the system's default personalisation* and 4% selected *Other*. Two who selected *Other* also provided the comments "to[o] much else to do!" and "didn't think it would make a difference".

The next question asked of users who *had* inspected the personalisation, what was their main reason: 42% selected *Curiosity*, 6% selected *I did not like how a page was personalised to me*, 6% selected *a belief about me was wrong* and 6% selected *Other*. Some who selected *Other* also provided comments:

- "Partly curiosity, partly because in the practice thing the sports field was irritating as it was not what I expected, and I didn't want to be irritated again". This user was commenting about their annoyance caused by the SASY Introduction Guide, that assumed users had an interest in sport and included additional content in the material due to this. The guide asked users to practice using the scrutinisation tools to determine why the content was included by SASY's personalisation.
- "Want to choose to do the thing I want to do". This comment is probably referring to the user having scrutinised to see whether they could skip content in the course.
- "No reason".

The final question of the survey allowed students to provide free form comments. The following were the responses:

- "I know all this stuff, and I should be able to skip to the post test without having to do all the other stuff". This was a fair comment since this user scored 9/10 in the pre-test, but SASY required users to complete the quiz for each module in addition to the pre and post tests.
- "I like the tick that appears when a question is answered correctly."
- "I did not like the color scheme". This is also a fair comment, one might expect in a commercial system that the user should be able to configure the colour scheme used by the interface.

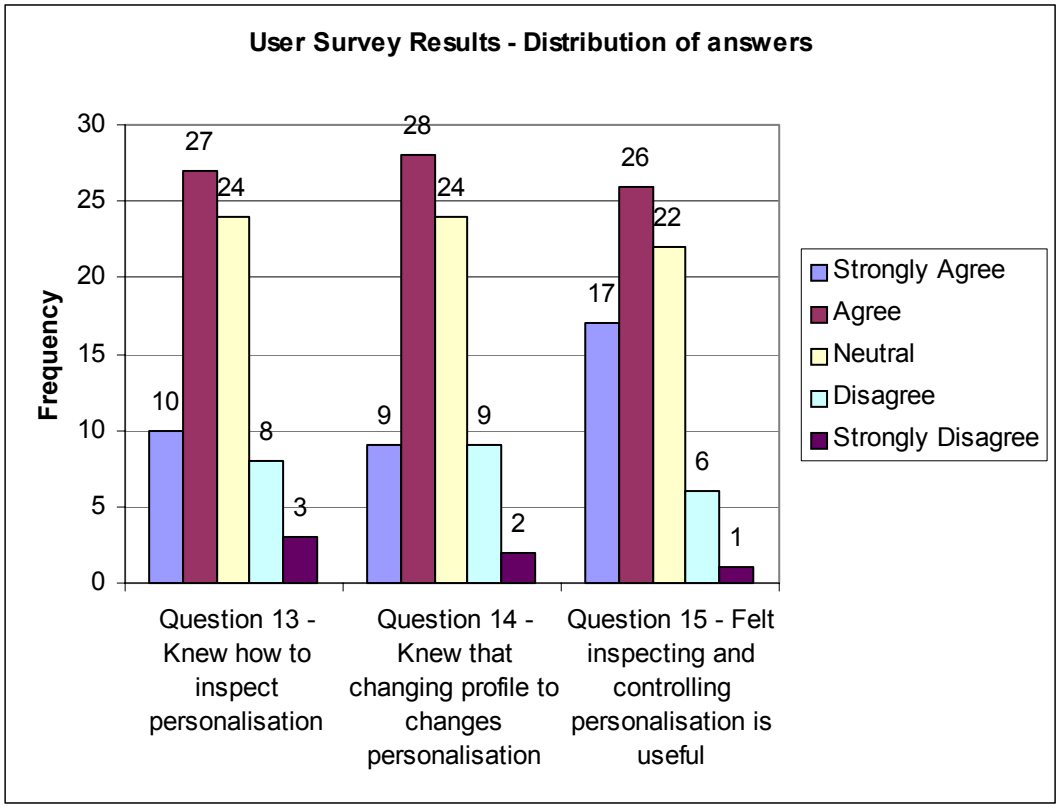


Figure 6.5 – Distribution of results to qualitative survey questions.

Additional Qualitative Feedback

In addition to the survey, we invited students by email to volunteer to further discuss their SASY experience. However, only one user (105) responded. This user was the highest user of the scrutinisation tools and had exhibited the following interesting behaviour:

- Sought evidence about their alleged lack of knowledge about a concept immediately after completing the pre-test, even though they demonstrated some level of knowledge in the pre-test by correctly answering some, but not all, questions about that concept. When asked about this, the user stated they were surprised SASY stated they did not know about UNIX Chmod and wanted to find out why it said this.
- Accessed evidence about their knowledge at the end of their SASY session, seemingly reflecting on their knowledge. When asked about this the user said "I did go back to see what I had learnt from SASY, and I was also curious about SASY and what it had determined about me. I also wanted to see if there was anything else I had missed. I chose to access this information at the end of the session largely because SASY gave feedback throughout the session as to what it knew about me, with the most useful information being which tests I had attempted and what I had yet to do. There was little need to see what the system had assumed about me until the end of the session, when curiosity took over."

The questions posed to this user and the responses are included in Appendix 8.4.10 verbatim.

6.2.5 Discussion

Expected forms of Scrutinisation

As described in the design, we had set up the personalisation so that by default it would add potentially undesirable content such as jokes, hints, many quiz questions and the presentation of advanced concepts. We expected this would motivate some users to scrutinise how a page was personalised, scrutinise their user model, scrutinise evidence of how a user model attribute was set, and/or change their user model so that the personalisation would:

A. Remove jokes and hints from the material.
B. Remove advanced concepts from the material.
C. Only show a small number of quiz questions included in the material.
D. Reflect the user had some knowledge about a concept though they did not answer every question about that concept correctly in the pre-test.

We also expected to see other, unprovoked forms of scrutinisation. We first discuss the expected forms of scrutinisation.

A. Scrutinisation of Jokes and Hints

Two (2%) users scrutinised evidence as to why SASY thought they wanted jokes in the teaching material and no users sought evidence about the Hints attribute. There were 4 different users (4%) who changed their profile to remove Hints, 8 (8%) removed Jokes. This result was surprising because we believed the jokes and hints were quite annoying and expected more users to exercise control over the personalisation to remove this content. However, in the qualitative survey, 93% of users said the jokes was either not annoying or not even noticeable. Similarly, 81% said the same thing about the hints.

B. Scrutinisation of Advanced Concepts

Five users (5%) scrutinised evidence as to why SASY thought they wanted to be presented with advanced concepts. Only one user changed their profile to remove advanced concepts.

C. Scrutinisation of Quiz Questions

Fifteen users (14%) scrutinised evidence as to why SASY thought they wanted to be presented with lots of quiz questions. Six (6%) users changed this attribute in their profile.

D. Scrutinisation of Knowledge

User model attributes that modelled knowledge about the domain were by far the biggest trigger for scrutinisation. In all, 37 (35%) users sought evidence about at least one of these three knowledge attributes. Of these, 11 (10%) sought evidence about their alleged lack of knowledge about a concept immediately after completing the pre-test, even though they demonstrated some level knowledge in the pre-test by correctly answering some, but not all, questions about that concept.

Also, 24 (22%) changed their knowledge profile attributes. Of these, 5 (5%) changed their profile to state they knew *more* about a module than SASY said they did, having just completed the pre-test answering most questions correctly.

Overall use of the Scrutinisation Tools

Since the user's main task was to learn about UNIX security, any scrutinisation was an aside activity and a distraction from their core task. In this context, overall scrutinisation was quite high: 81 users (77%) scrutinised in some way, 13 (12%) scrutinised 9 or more times (i.e. more than twice average). However, the usage of each tool separately is not so high since users on average used 2 forms of scrutinisation. There was a correlation (coefficient=0.76) between the number of forms of scrutinisation and the total scrutiny actions, meaning that users who scrutinised more overall tended to use more forms of scrutinisation. The average number of scrutinisation actions was 3.
Usage of all the SASY scrutinisation was observed in the logs. The following list is a breakdown of the number of users who scrutinised with each of the SASY scrutinisation tools:

- Highlight Tool – 36 (34%) users inspected which aspects of the hypertext were personalised (i.e. added to and/or removed) and which aspects of the user model caused each personalisation. In the qualitative survey, only 11 users (15%) strongly disagreed or disagreed they knew how to inspect the personalisation to see what items were included or removed from their personalised view. There were 7 (7%) users that used the Highlight tool to access the pre and/or post test when the personalisation changed to no longer display a link to it. One concern with adaptive hypertext is that as it changes over time, users can not access the view of the page they had seen previously. This observation shows users used the Highlight tool to do just that.
- Evidence Tool – 45 (43%) users viewed evidence of how user model attributes were set, whether set directly by the user or inferred by the system.
- Profile Tool – 49 (47%) users viewed their profile and 37 (35%) changed aspects of their user model to control the personalisation. In the qualitative survey, only 11 users (15%) strongly disagreed or disagreed they knew that they could change the beliefs on the profile page to change what was included and removed by the personalisation.

Overall, the usage statistics indicate users were able to find, access and use the scrutinisation tools. Although the tools are may not be used often or heavily, when a user wants to scrutinise, they should be able to easily find and use the tools. The qualitative survey questions (13 and 14) asked users whether they knew about the tools. Since there were only 11 users (15%) to these questions, this indicates a large group were unsure whether they really did know how to use the tools. As discussed, there was some inconsistency in the responses the qualitative survey questions surrounding whether users knew how to change the personalisation to remove hints and jokes.

Timing of scrutinisation

From the logs we observed most scrutinisation actions (96%) occurred when the students were either about to start, or had just finished a learning module, session, pre or post test. All the scrutinisation tools were used at these times. In addition, the course map was mostly used at this time. Therefore, not many students scrutinised while they were learning but rather just before or after learning.

The Evidence and Highlight tools were mostly invoked from the home page. This is understandable since the home page was the navigational hub of the course. Every session started on this page, every module started and ended on this page. Also, as users completed a module successfully, the navigation in SASY directed them back to the home page which would then look different compared to the last time it was accessed, with navigation changed elements now directing the user to a different module. Perhaps this change motivated users to scrutinise to either understand why the page had changed, to attempt to return the page to the previous state, or access links that were available previously but had been removed by the personalisation. For example, there were 7 users who attempted to access the pre or post test multiple times.

Control over personalisation

In addition to the 37 (35%) students who changed aspects of their user model to control the personalisation, there were a few interesting observations about the degree to which users exercised control over personalisation.

Firstly, 10% of all scrutinisation actions occurred before users completed the pre-test. This indicates not only that the scrutinisation tools were noticeable immediately, but also that some users were quite keen to be involved in the personalisation process before they had even accessed the course material, though they may have done another topic first.

Some users demonstrated understanding over the personalisation and how they could steer it to meet their needs. For example, two users changed their profile to state they had less knowledge, having completed the whole course once already. These users seemed to be revising the course or wanted to repeat the learning experience.

Seven (7%) users changed the system preference in their profile to highlight personalisation on every page by default. Of these, 3 users worked in this mode for the whole course or a substantial part of it. Feedback from one user who did this indicated they were interested in how the system worked and personalised the material to them.

Finally, in the qualitative survey, 43 users (60%) strongly agreed or agreed it was useful to be able to inspect and control the personalisation, and just 6 users (8%) disagreed and one strongly disagreed.

6.3 Experiment 2 – Detailed Laboratory Study

6.3.1 Design

This study was qualitative, complementing Evaluation 1. Its main purpose to capture data about the usability of the tools and use this information to enhance them. The end goal was a set of tools that allow the user to easily and effectively scrutinise an adaptive hypertext in the following ways (**Goal 1**):

A) Understand what aspects of the hypertext were personalised. That is, what parts of the content were added to and/or removed from the user's personalised view of the page.
B) Understand what aspects of the user model caused each personalisation. See the implications of changing their user model, in terms of how it effects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes.
C) View and understand evidence of how user model attributes were set, whether set directly by the user or inferred by the system.
D) View or change aspects of their user model to control the personalisation.

The second goal of this evaluation (**Goal 3**), was to obtain feedback about the importance of scrutability and control of personalisation from a broader audience than that of Evaluation 1.

Finally, as per **Goal 4**, we intended to evaluate the scrutability tools in a different domain and in a different type of personalisation system, than Evaluation 1. This study evaluates SASY when used as a personalised recommender system, as opposed to a personalised teaching system in Evaluation 1.

To satisfy these evaluation goals, we designed this experiment to have participants to perform a series of tasks to scrutinise a Personalised TV Guide. SASY recommended a personalised television program viewing schedule to users, based on their interests and preferences defined in their user model. User interaction was structured into small tasks, so that users would provide feedback between tasks by answering short, targeted questions. We base this approach on Paramythis et. al. (2001), who argue that to evaluate adaptive systems one needs to understand how users experience and perceive the user model during the interaction. Users were not directly asked to use the scrutinisation tools but had to figure out how to access and use the tools in order to complete the tasks. The tasks covered all the forms of scrutinisation listed above in items A through D. The ability of users to use the tools to complete the tasks, without training or help, would indicate the *effectiveness* of the scrutinisation tools.

At the end of their session, users completed an online questionnaire and additional feedback was obtained through informal discussion with participants following the evaluation. The purpose of the questionnaire and discussion was to measure user satisfaction in regards to use of the tools and their opinion on the importance of scrutability and control of personalisation.

Initially, we expected to conduct this study once to prove the scrutinisation tools in SASY 1 were effective. However, following the pilot study with SASY 1, it was clear that minor enhancements to the user interface could give valuable improvements to user performance on tasks. Therefore, we conducted four further iterations of the evaluation, as described in Table 6.16. The results of each iteration were used to refine the scrutinisation tools to create a new system variant, that was used for the next iteration. In iterations 1 and 2, users were required to work through brief online training. No training was provided in the other iterations. The differences between the system variants are discussed in the following sections where relevant.

Table 6.16 – Summary of Iterations of Evaluation 2.

Iteration	Aim & Method	System
1	**SASY 1 (Pilot Study) with online training**. We first conducted the experiment using SASY 1, the same version as used in Evaluation 1, to obtain feedback about the effectiveness of the scrutinisation tools in SASY 1. We required participants to read the online training, which was the SASY Introduction Guide from Evaluation 1 since we wanted to evaluate this material to determine whether it was helpful. Participants then completed the worksheet tasks to scrutinise a Personalised TV Guide.	SASY 1
2	**SASY 2 with online training** We used results from Iteration 1 to refine the interface to the scrutinisation tools to create SASY 2. The aim of this iteration was to evaluate the scrutinisation tools in SASY 2. Method was the same as for Iteration 1.	SASY 2
3	**SASY 2** Aim was to evaluate the effectiveness of SASY 2 scrutinisation tools on untrained users. We repeated Iteration 2 but this time with no online training, to determine whether this would impact task performance.	SASY 2
4	**SASY 3 – Added Just-In-Time Help** In previous iterations, users had difficulty understanding what caused each adaptation using the Highlight tool (scrutinisation B). We added just-in-time help for the Highlight tool to create SASY 3, and repeated the experiment to determine whether his would improve task performance with the Highlight tool. Method same as Iteration 3.	SASY 3
5	**SASY 4 – Created different version of Highlight Tool** Iteration 4 did not solve the problem with the Highlight. We created a variant of the Highlight tool, SASY 4, to reduce mouse interaction. The purpose of this iteration was to evaluate the Highlight tool in SASY 4 to determine if this variant would improve task performance with the Highlight tool. Method same as Iteration 3.	SASY 4

6.3.2 Materials

Personalised TV Guide

The TV Guide content was just a regular topic in SASY, developed using the same ATML framework as Evaluation 1. On logging in to the TV Guide for the first time, the initial profile page required the user to specify whether or not they were interested in the following program genres: Sports, Business & Finance, Current Affairs, Drama, Children's programs and Lifestyle programs. Following the initial profile, the TV Guide content was presented. This consisted of two pages of personalised TV Programs schedules, where programs were added or removed from the user's personalised schedule based on whether the programs matched the user's interests as defined by their profile (user model). For example, Figure 6.6 shows the first TV program schedule page as it appears in SASY 1. The page shows the recommended programs, grouped by timeslots 4:30am – 5am and 5am-5:30 am. The right hand side of the page summarises the personalisation that occurred on the page. Figure 6.6 shows 2 items were removed from the page, 4 items were added. The right had side also lists the attributes from the profile that affected personalisation. The first bullet point in the figure states the user is not a member of *SIG 1*, a special interest group for religious themes. Note that it is not obvious what SIG 1 means by looking at this page alone, as was intended by the evaluation design so that users could not guess its meaning without using the scrutinisation tools. This was intentional and is explained later. The TV Guide also included a single a Program Index page, where the user could click on a TV program to pop-up a description about it.

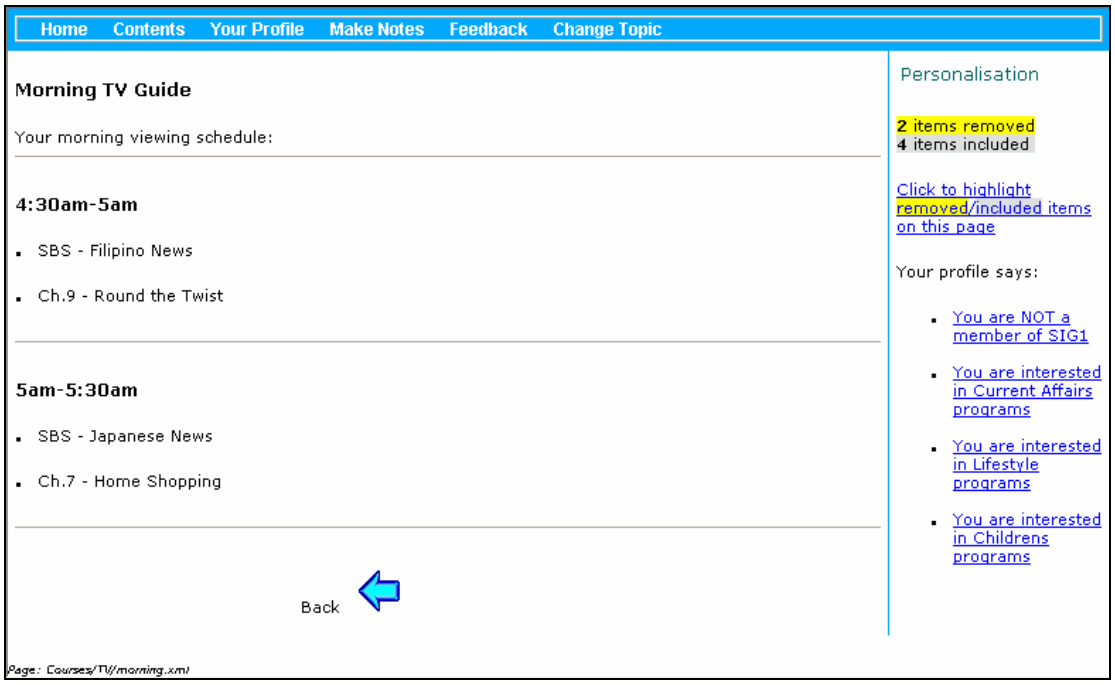


Figure 6.6 – Sample page from the Personalised TV Guide Evaluation in SASY 1.

The profile page also contained four attributes that were initially hidden and set by defaults, but could be changed by the user if they accessed the profile after the initial profile creation. The first specified whether the user had a channel preference, with possible options being: Ch. 2, Ch. 7, Ch.9, Ch. 10, SBS and None. By default *None* was selected indicating the user had no preference so they were shown programs from all channels. If a preference was selected by the user, only programs playing on the selected channel were added to their personalised TV Guide, all other programs were removed. The other three, initially hidden profile attributes, indicated whether or not the user belonged to Special Interest Groups (SIGs). The groups were intentionally obscurely named as SIG 1, SIG 2 and SIG 3 so that users would not be able to guess their meaning to complete tasks. SIG 1 was a special interest group for people interested in programs with religious themes. SIG 2: health issues. SIG 3: crime. By default users were not members of any group, meaning they were not shown programs that belonged to these genres. However, if the user accessed the Program Index page and clicked *any* program to see a description about it, SASY made them a member of SIG 1 (religious programs) regardless of what program was clicked. This was intentional as we had set tasks (Tasks 9 and 10 in Table 6.17) that required the user to click any program, then return to their personalised schedule and investigate, using the scrutinisation tools, why their profile had been changed and how this effected personalisation.

Worksheet Tasks

Tasks were carefully designed to test whether participants are able to learn and use the scrutinisation tools in without prior training, since this is how people will need to be able to use the tools in practice. Participants were asked to complete an online worksheet, that presented one task at a time. On completing a task, participants had to click a *Continue* button to be shown the next task. Each task required users to interact with their personalised TV Guide in SASY, but did not instruct users to use the scrutinisation tools. Nor did it explain which tool to use or how to use the tools. To complete tasks, user's had to figure out they needed to use the scrutinisation tools, explore the interface to determine which tools to use and how to use them.

The first column of Table 6.17 has the actual worksheet tasks as shown to users in the evaluation. The other columns in the table were not shown to users but are provided here to describe the purpose of each task, the form(s) of scrutinisation that was required to complete the task, a reference to the scrutinisation form(s) listed in the design of this evaluation (A, B, C and D). For example, Task 2 required participants to examine the personalised TV Guide page, as in Figure 6.6, and find out all the programs that play in this timeslot. To complete this task, participants had to find and access the Highlight tool, scrutinisation form A, to make visible all programs that had been removed from the personalised view. For Task 2 only, a *Hint* hyperlink was provided that, if clicked, popped up the help text "In SASY, look at the right hand side of the page in the personalisation column".

The personalisation in Tasks 2 to 6 was based only on their initial profile preferences. Tasks 7 through 11, required participants to appreciate personalisation was not only based on their initial preferences, but also attributes that had been inferred about them by SASY. Task 8 required the user to scrutinise

nested adaptation rules. By our definition, as in Chapter 1, Tasks 2-6 represented simple forms of personalisation while Tasks 7-11 were ore complex.

The design of the tasks was such that to successfully complete them, users would have to:

- use all the scrutinisation tools: Profile tool, Highlight tool and Evidence tool;
- understand personalisation was based on their initial preferences and attributes that had been inferred about them by SASY as they used the system;
- exercise control over the personalisation by changing their profile.

Table 6.17 – Description of worksheet provided to participants.

	Task	**Purpose / What is being evaluated**	**Forms of scrutinisation needed to complete task**
1	Login and fill in initial profile questionnaire freely.	Populate the user model with the user's interests in TV program genres (e.g. Sport).	N/A
2	• Examine personalised TV Guide page. • Find out all the programs that play in this timeslot.	Ability to scrutinise to see *what* aspects of a page were adapted using the Highlight tool.	A
3 to 5	Which programs are added for people who are interested in • Current Affairs programs? • Children's programs? • Lifestyle programs?	Ability to scrutinise to understand *why* each adaptation occurred. We expected users would use the Highlight tool (A, B) to access explanations of each personalisation. It was also possible, albeit slower, to complete these tasks by directly changing the profile (D) and observing the effect on personalisation.	A and B *or* D
6	Change the personalisation so that only Current Affairs programs are included in your 4:30 – 5:30 schedule.	Ability to control personalisation by changing profile.	D

	Task	**Purpose / What is being evaluated**	**Forms of scrutinisation needed to complete task**
7	Change the personalisation so that the program **Aerobics Oz Style** is included in your schedule.	Ability to understand why personalisation occurred and how to exercise control over it, where personalisation occurred due to an initially hidden user model attribute. Here personalisation occurred because they are not a member of SIG 2 – special interest group for sports programs.	A and B and D *or* D
8	Change the personalisation so that the current affairs program that is included in your 6 – 7am schedule is the **Today** show.	Ability to understand complex form of personalisation where the personalisation is based on *nested rules*, involving initial and inferred user model attributes.	A and B and D *or* D
9	Access Program Index page and freely click on programs of interest.	SASY changes the user's profile in the background to make them a member of a SIG 1 – a special interest group for people who are interested in programs with a religious theme.	N/A
10	Return to TV Guide constructed earlier and explain: • Why it has changed? • Why programs have been added? • Why system states you are a member of a special interest group interested in religious themes?	Since SASY changed the participant's profile in Task 9, additional programs now appear in the TV Guide because the participant is now a member of SIG 1. Task evaluates the use of the Highlight and Evidence tools. Additionally evaluate whether the participant appreciates their profile has been changed by SASY without their consent and how they feel about this.	A and B and C
11	Change the personalisation such that only Current Affairs programs are included in your 4:30am – 5:30am schedule, as before.	This is a repeat of Task 6 to see if the participant can perform this task more easily than previously.	D

Qualitative Survey

Following the completion of the worksheet, participants were asked several qualitative questions to get their feedback about their experience using the scrutinisation tools. The purpose of the survey was to gauge whether participants believed they understood how to use the scrutinisation tools and measure their satisfaction and acceptability of the tools. The survey questions are presented in Table 6.18. For each, the user was asked to select one of the possible answers: Strongly Agree, Agree, Neutral, Disagree and Strongly Disagree and optionally provide free form comments.

Table 6.18 – Description of survey following worksheet completion for Evaluation 2.

	Question	Purpose / What is being evaluated
1	I knew how to inspect the personalisation to see what items were included or removed from my personalised view.	Determine whether participants knew about the highlight tool.
2	It was easy to inspect the personalisation to see what items were included or removed from my personalised view.	Determine whether participants found the highlight tool easy to use.
3	I knew that by changing my answers to questions on the Profile page I could change what was included and removed by the personalisation.	Determine whether participants knew about the profile tool.
4	It was easy to change what was included and removed by the personalisation.	Determine whether participants found the profile tool easy to use.
5	I feel it is useful to be able to inspect and control personalisation.	General question to gauge how participants felt about scrutability and control over personalisation.
6	Overall, I felt in control on the personalisation.	Gauge whether participants felt in control of the personalisation.
7	Overall, I understood how the personalisation worked.	Gauge whether participants understood how the scrutinisation tools worked.

6.3.3 Participants and Procedure

We selected participants from personal and professional acquaintances. The group was selected to represent potential users of the scrutinisation tools. As such, participants had ranged in computer literacy, socio-economical background, career, age and gender. In all, there were thirteen participants. Table 6.19 summarises participant backgrounds. All had tertiary education qualifications in relation to their current career, except P3 who had studied tourism. All were adults. All were familiar with internet browsers. Most of the participants had a medium level of computer literacy, except P2 and P3.

Table 6.19 – Evaluation 2 Participants.

Computer Literacy codes: C = Computer application development, I = Internet browsing, IB = Internet banking, IM = Downloading music through Internet, R = Records Management, S = System Administration, W = Word processing

Participant	Age	Gender	Career/Profession	Level of computer literacy	Computer usage Hours / week	Iteration
P1	20-30	Female	Marketing	Medium (W, I, IB)	35	1
P2	30-40	Male	Chef	Low (I)	1	1
P3	30-40	Female	Home maker	Low (W, I)	3	1
P4	40-50	Male	Business Analyst	High (W, I, IB)	40	2
P5	30-40	Male	Business Analyst	Very High (R, W, I)	40	2
P6	20-30	Male	IT Developer	Very High (C, I)	40	3
P7	30-40	Male	Builder	Medium (I, IM)	10	3
P8	20-30	Female	Tax	Medium (W, I, IB)	40	3
P9	20-30	Male	Tax	Medium (W, I, IB)	40	3
P10	30-40	Male	IT Support	Very High (S, I)	40	4
P11	30-40	Female	IT Support	Very High (C, I)	40	5
P12	30-40	Male	IT Developer	Very High (C, I)	40	5
P13	30-40	Male	Restaurant Manager	Medium (I, IB)	5	5

Different participants were involved in each iteration. The sample size in each iteration was small, based on Nielsen (1994), who argues that three to five participants are adequate for usability studies, since evaluation with five participants will uncover about 85% of potential usability problems. In addition, evaluation with one participant will already uncover about a third of usability problems, with a diminishing amount of information discovered with each subsequent participant. For this reason, the number of participants recruited for each iteration varied. We evaluated with as many participants as we deemed necessary and stopped when we believed we had enough information to draw conclusions. Where we identified a problem, we ended that iteration, and addressed the problem before moving to the next iteration.

Participants took part in the experiment individually during separate sessions. They completed a worksheet that described tasks they had to perform to create a personalised TV Guide in SASY. Participants were asked to express any points of confusion, but to avoid cognitive load, were not required to verbalise their thought process while completing tasks. Participants were observed, comments and actions were noted and their actions were logged by the system in the background. Following evaluation, users completed an online qualitative survey. Then we informally discussed the

evaluation and scrutinisation tools with them to capture any additional comments. There was no other interaction with participants unless they could not continue for some reason. All interventions are discussed in the results.

6.3.4 Iteration 1 – SASY 1 (Pilot Study) with online training

We first conducted the experiment using SASY 1, the same version as used in Evaluation 1, to obtain feedback about the effectiveness of the scrutinisation tools in SASY 1. We required participants to read the online training, which was the SASY Introduction Guide from Evaluation 1 since we wanted to evaluate this material to determine whether it was helpful.

SASY 1 overview

Since the SASY 1 interface differs slightly to SASY 4 (presented in Chapter 3), it is briefly discussed here. The Profile and Evidence tools are the same as in SASY 4. The key differences are elements on the right hand column of a personalised page and the Highlight tool.

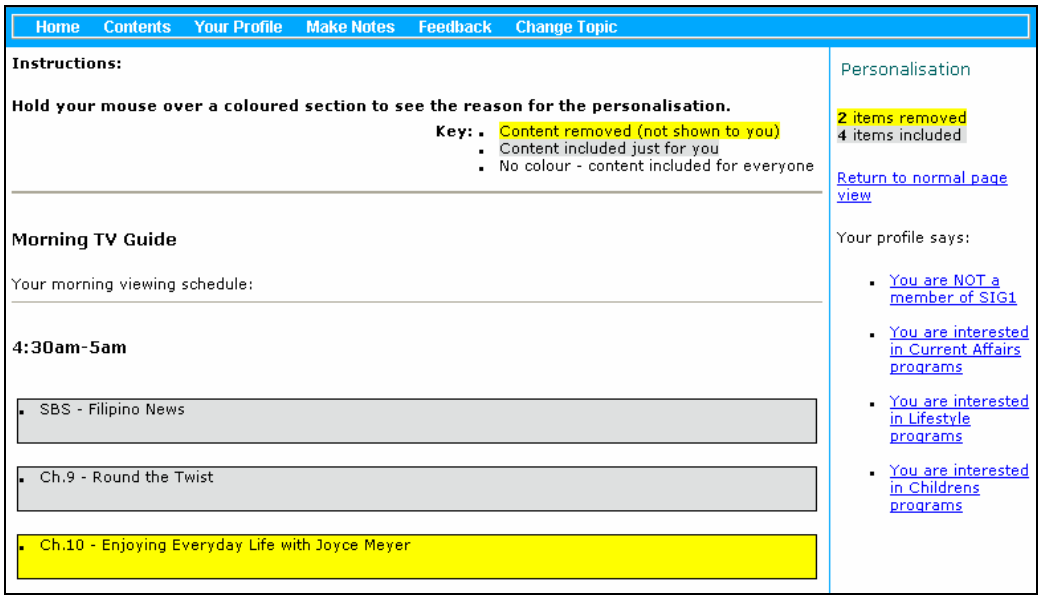


Figure 6.7 – Highlight tool in SASY 1. Invoked from personalised page in Figure 6.6 (page 180).

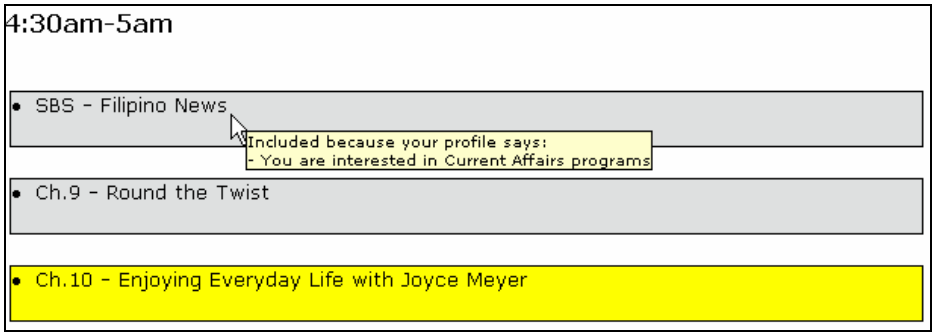


Figure 6.8 – Mouse over explanation of Highlight tool in SASY 1. Explanation message was popped up by moving mouse pointer over the grey area "SBS – Filipino News".

Figure 6.6 (page 180) shows a screen shot of a personalised page in SASY 1. There is an additional link at the top right of the page, "Click to highlight removed/included items on this page", that also invokes the Highlight tool. The Highlight tool is also different: SASY 1 only has the Mouse-Over mode of the tool. In this mode, the Highlight tool reloads the page to annotate with colour all the personalised content, as shown in Figure 6.7. However, there is no explanation text initially provided as to why each item was added or removed by personalisation. To view this explanation, the user has to move their mouse pointer over a highlighted area. This pops up the appropriate explanation text. For example, moving the mouse cursor over the first grey area in Figure 6.7, over the text "SBS – Filipino News", pops up the message "Included because your profile says you are interested in Current Affairs programs", as shown in Figure 6.8. When the Highlight tool is invoked, the top of the page contains a key that describes the colour coding used to highlight content that was removed and added to the personalised view. It also instructs the user to move their mouse over a highlighted area to see the reason for the personalisation.

Iteration 1 – Results

There were three participants in this study: P1, P2 and P3 (refer to Table 6.19). Appendix Table 8.15 (page 250) shows the forms of scrutinisation participants explored in completing the tasks. It compares the expected forms of scrutinisation with the results. The table shows that users in this pilot iteration had considerable difficulty completing the tasks with ? indicating the tasks that users could not complete. We discuss the results in terms of the SASY scrutinisation tools and the forms of scrutinisation they are designed to support. Observation notes of each participant are in Appendix 8.5.2 (page 254).

Online help

All three participants read through the online help material, the same material that was voluntarily accessed by some user in Evaluation 1. Only P1 earnestly practiced scrutinising which reading the guide. The other users read the material but did not practice. All three participants commented the material was overly verbose and complicated.

Highlight Tool (Scrutinisation A and B)

Only P2 found and accessed the Highlight tool without help and used it to determine what was adapted on a page (scrutinisation form A). P1 required the use of the Hint in the worksheet to find the tool, P2 did not use tool at all.

In order to understand why each content item was personalised users needed to access the mouse over explanations using the Highlight tool (scrutinisation form B). None of the participants noticed this feature, so could not effectively use the tool beyond Task 2.

For Task 7, P1 requested help, at which point the instructor asked P1 to read the instructions at the top of the page when the Highlight tool was invoked. P1 read the instructions and then understood how to use the tool. P1 then used of the tool in Task 8 without help. Likewise P2 and P3 were directed to read the instructions at Task 10, which allowed them to complete the task.

Feedback from the informal discussion with users following the evaluation indicated it was not clear how to access the Highlight tool from the interface, for two reasons. Firstly, in SASY 1, the *N items added* and *N items removed* elements did not look like hyperlinks since they were not underlined like other hyperlinks. Secondly, participants did not understand the purpose of the link "Click to highlight removed/included items on this page". The users commented the label was confusing.

Evidence Tool (Scrutinisation C)

None of the participants were initially able to find the Evidence tool. In Task 10, the instructor asked participants to carefully examine the interface, after which all eventually found the links that accessed the evidence. Once participants accessed the tool they acknowledged they understood its purpose.

Profile Tool (Scrutinisation D)

Users easily grasped the concept that the profile was a model of their interests and preferences and controlled personalisation. Also users understood that they could change their profile to change what was added and removed from the schedule to create their personalised view. All participants quickly and easily completed Tasks 6 and 11 that required them to change their profile directly.

Qualitative Survey

Appendix Table 8.16 (page 252) shows participant responses to the qualitative survey questions. Note that in this first iteration, the survey did not include Questions 2, 4, 6 and 7 (so these are blank in the table). For all responses, users answered they agreed or strongly agreed they knew how to inspect personalisation and change their profile. P2 and P3 commented they did not initially know but once they worked it out it was quite evident. They all agreed it was useful to be able to inspect and control personalisation.

Iteration 1 – Summary

Despite having read the SASY Introduction Guide that briefly explained the scrutinisation tools, it was not obvious to participants how to access the Highlight and Evidence tools in SASY 1. Users who did access the Highlight tool did not notice the mouse-over explanations. Users seemed to understand the purpose of these tools once they accessed it, but could not initially find them. On the other hand, users had no difficulty using the Profile tool. These user interface issues caused confusion for the users, whose poor performance made it difficult to evaluate the true effectiveness of the tools.

We refined SASY 1 to create SASY 2 that addressed the issues expressed out the Highlight tool. We also refined the wording of tasks in the worksheet.

6.3.5 Iteration 2 – SASY 2 with online training

We used results from Iteration 1 to refine the interface to the scrutinisation tools to create SASY 2. We also refined the online training material. We used the same method as Iteration 1, with two participants: P4 and P5 (refer to Table 6.19).

SASY 2 overview

Figure 6.9 shows a screenshot of SASY 2. There were two, subtle yet important, enhancements made from SASY 1. Firstly, in the right hand side of any personalised page, the items removed and items added links were given the same font as other hyperlinks so they could be immediately recognised as links. We noticed in Iteration 1 that users understood clicking these links would show (in some form) what the removed and added items were. Therefore, we removed the link "Click to highlight removed/included items on this page" as it was confusing and redundant.

Secondly, we attempted to make the Highlight tool instructions and mouse over explanations stand out more. Figure 6.10 shows a screen shot of the Highlight Tool invoked in SASY 2. The instructions at the top of the page are in a larger font than in SASY 1 and encapsulated by a box. Also the mouse over explanations were now in a larger font, with the first word, "Added" or "Removed", displayed in bold. The background colour of the mouse over message was changed to be a slight offset of the colour of the highlighted text that it related to. For example, in Figure 6.10, the mouse cursor is hovered over the first grey area with text "SBS – Filipino News". A slightly lighter shade of grey was used for the pop-up message "Added because profile has…". Likewise, for removed items, a lighter shade of yellow was used for the pop up message. The purpose of this was to distinguish the pop-up messages from regular tool tip messages used by other HTML elements, with the aim of making the pop-up messages more noticeable. We also made the SASY Introduction Guide more concise, to make it easier to read. We placed a radio button at the bottom of every page in the Guide, requiring users to indicate whether or not they understood the information on that page.

Iteration 2 – Results

Appendix Table 8.15 (page 250) shows the forms of scrutinisation participants explored in completing the tasks, observation notes are provided in Appendix 8.5.3 (page 256).

Online help

Both participants read through the online help material and on all training pages, answering Yes to the question "Do you feel confident you understand the information above?". Both practiced using the scrutinisation tools and appeared to understand them – an improvement over Iteration 1.

Highlight Tool (Scrutinisation A and B)

Both easily found and accessed the Highlight tool. Both initially forgot how to access the mouse over messages for the Highlight tool, but after a minute found the instructions and were able to use the tool effectively for scrutinisation forms A and B. Like participants in Iteration 1, for Task 7, P4 first tried to directly change the profile, but when they realised they could not work out the answer this way, P4 returned to the personalised page and used the Highlight tool to complete the task as expected. P5 had no difficulty using the tool but stated "I'd forgotten about the pop up. It's something you just have to learn about or get reminded."

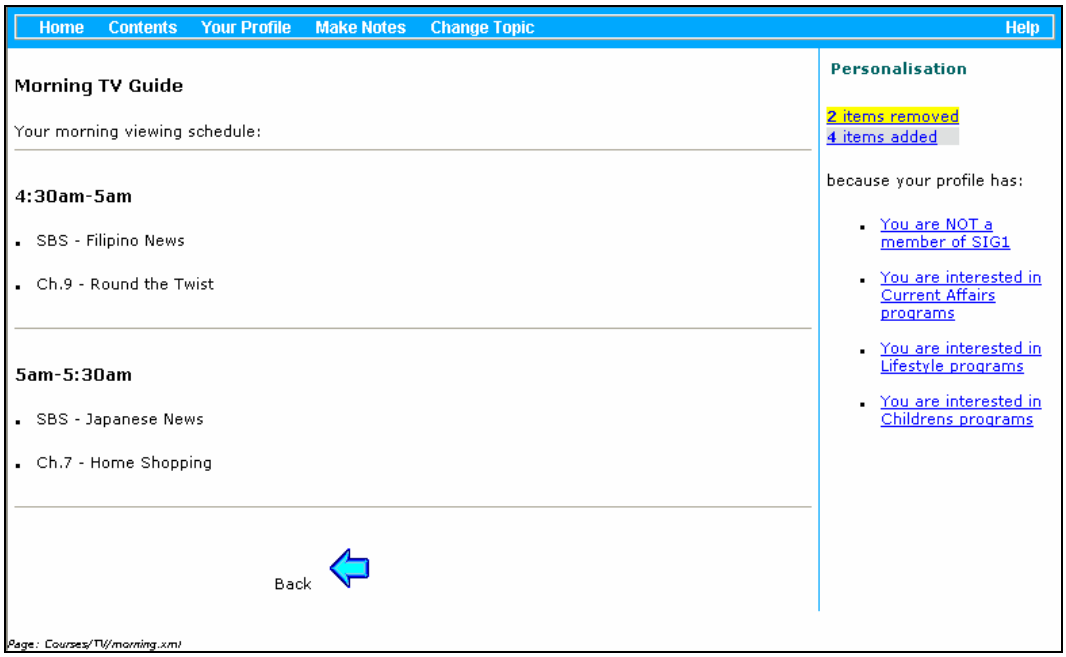


Figure 6.9 – Screenshot of SASY 2.

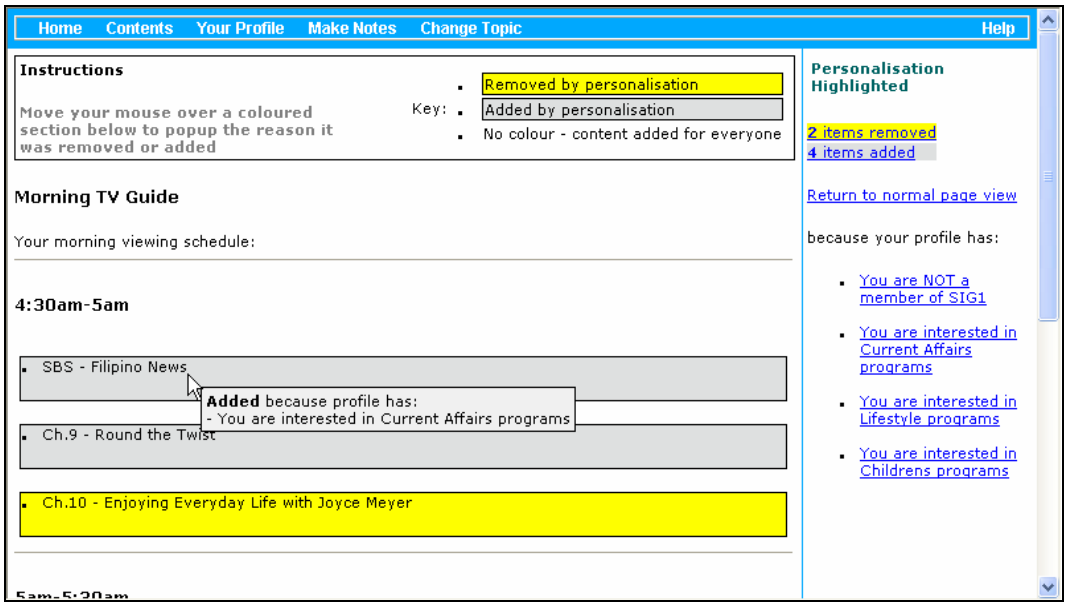


Figure 6.10 – Screenshot of SASY 2 with Highlight tool invoked.

Evidence Tool (Scrutinisation C)

P4 accessed the Evidence tool multiple times, mostly as a means of accessing the profile. However, in Task 10, did not think to use the Evidence tool to complete the task. P5 did not initially know how to use the Evidence Tool but found it after exploring the interface. Following the evaluation, P4 suggested adding the label "Why?" for the evidence tool, similar to how www.amazon.com labels its links that allow users to seek explanation for personalised content.

Profile Tool (Scrutinisation D)

P4 did not notice the link to the profile in the top menu. Instead, P4 continuously accessed the profile by first accessing the Evidence tool and clicking though to the Profile tool. As in Iteration 1, both users were comfortable using the Profile tool.

Qualitative Survey

For all responses to the survey (Appendix Table 8.16, page 252), participants answered they agreed or strongly agreed they knew how to inspect personalisation and change their profile. Both agreed it was useful to be able to inspect and control personalisation.

Iteration 2 – Summary

Participants of Iteration 2 performed better than participants in Iteration 1. Participants easily found and accessed the Highlight tool (scrutinisation form A) and Profile tool (scrutinisation form B), to scrutinise simple personalisations.

However, there were two potential problems with the interface. Firstly, users still could not easily access the pop-up explanations of the Highlight tool (scrutinisation form C). Secondly, users did also did not seem to understand the purpose links that invoked the Evidence tool (scrutinisation form D). However, users managed to find both tools after examining the interface carefully.

Since both the online training material and the SASY user interface had changed since the previous iteration, it could not be determined which lead to the improved results. Mostly likely, both played a role. We decided to stop evaluating after two participants, and repeat the experiment for additional participants without training, to determine whether untrained users would still be able to find the tools.

6.3.6 Iteration 3 – SASY 2

The purpose of this iteration was to evaluate SASY 2 without online training, to determine whether the interface was usable on first contact. We used the same method as Iteration 1 but this time users did not read the training material. There were four participants in this study: P6, P7, P8 and P9 (refer to Table 6.19).

We made one change to the interface following Iteration 2, changing the appearance of the link(s) that accessed the Evidence tool. Following the suggestion of participant 4 in Iteration 2, we added the label "Why?" to the link label. Participant 4 was a user of www.amazon.com (this website uses a "Why?" label in its interface to allow users to seek evidence of recommended items) and claimed this was more intuitive than the SASY interface.

Iteration 3 – Results

Appendix Table 8.15 (page 250) shows the forms of scrutinisation participants explored in completing the tasks, observation notes are provided in Appendix 8.5.4 (page 259).

Highlight Tool (Scrutinisation A and B)

All, except P6, quickly found the Highlight tool (scrutinisation form A). This was a significant improvement over SASY 1. P6 was stuck at Task 2 and clicked a hyperlink in the worksheet to reveal the hint "Look at the right hand side of the page". P6 then found and invoked the Highlight tool.

However, having accessed the Highlight tool, only P6 noticed the mouse-over explanations (scrutinisation form B). The others did not notice them nor did they read the instructions at the top of the page that instructed them to move the mouse over highlighted items. Even P6 did not initially notice the instructions and only did so upon close inspection of the interface of the Highlight tool.

Once prompted by the instructor to read the instructions for using the mouse-over tool at the top of the page in SASY, P7, P8 and P9 were able to access explanations. Once participants found the mouse over explanations, the explanation text was sufficient to explain why personalisation had added or removed the content.

Evidence Tool (Scrutinisation C)

P6, P7 and P8 used the tool initially for navigation to the profile page. P6 and P9 quickly accessed the evidence link when looking for an explanation why a profile attribute had the current value. P7 and P8 first tried looking in the profile, but later found the evidence link also. The results suggest that changing the Evidence link to include the label "Why?" had helped users find and access the tool.

Profile Tool (Scrutinisation D)

As in previous iterations, users were confident in viewing and changing their profile. P6, P7 and P8 did not initially see the link to the profile in the menu. This was because the menu is not visible once the user scrolls down the web browser page.

Qualitative Survey

As shown in Table 8.16, responses were mixed. Participants commented that it was not obvious how to use the Highlight and access the mouse over explanations initially. P6 remarked "Being able to hover over an item and see what 'rules' were applied was useful." P9 commented "Pop up boxes assisted in modification of personalisation preferences. Once I understood how the system could be manipulated I was able to gain more control over the programs I wanted to watch." P8 commented "Initially I wasn't totally confident on what I was doing, but after a couple of attempts I could navigate around the system easily." After discussion with the participant, it was noted that they were referring to the use of the profile, highlight/mouse-over tools. P7 felt the Highlight tool, in particular the mouse-over pop-up, was difficult to use and commented "It was not obvious [what was added/removed]. I needed to fiddle around to find the answers."

All participants agreed or strongly agreed they knew that by changing their answers to questions on the Profile page they could change what was included and removed by the personalisation. However, P7 commented "Not always because I did not know what SIG 1 and SIG 2 meant." Most felt the profile was easy to use but P8 commented it might be better to be able to change the profile from the TV schedule page rather than having to navigate to the profile page.

All participants agreed or strongly agreed they feel it is useful to be able to inspect and control personalisation. P6 added "Especially true for being able to change the automatic special interest group allocation".

All participants agreed or strongly agreed they felt in control on the personalisation and understood how the personalisation worked. However, P9 commented "I think that I knew more about the special interest group setup this would have assisted in getting more control over the programs I wanted to watch. Once I understood how the system could be manipulated I was able to gain more control over the programs I wanted to watch."

Iteration 3 – Summary

Comparing the result from this experiment to Iteration 2, it is evident that the SASY Introduction Guide (online training) did help the users to understand and better use the scrutinisation tools in Iteration 2. In particular, having practiced accessing the mouse over explanations with the Highlight during training, improved performance during tasks. Participants in this Iteration, unlike the trained participants in Iteration 2, were not able to easily find the explanations (scrutinisation form B).

Whilst the results were not as good as for Iteration 2, it was definitely an improvement over Iteration 1. Using SASY 2, participates were able to quickly find the Highlight tool, and most found the Evidence tool, whereas participants could not do this using SASY 1. The interface changes pertaining to these tools had improved task performance.

The most significant problem was that once users accessed the Highlight tool, users still did not notice the mouse over explanations. This discouraged further use of the tool. The next step was to increase the visibility of the mouse over explanations, to better support scrutinisation form B.

6.3.7 Iteration 4 – SASY 3

The main problem with SASY 2 was that users could not access the mouse over explanations when using the Highlight tool. These explanations were a key part of the tool since they explained why content was added or removed to create the personalised view.

We made a system variant, SASY 3, that popped up another browser window containing instructions on how to access the mouse-over explanations when the user accessed the Highlight tool for the first time. The pop-up is shown in Figure 6.11. We expected users would easily notice the instructions in the pop-up and would be able to access the mouse over explanations.

The purpose of this Iteration was to evaluate whether the just-in-time help window helped users understand how to use the Highlight tool, for scrutinisation form B. We repeated the experiment with only one participant (P10).

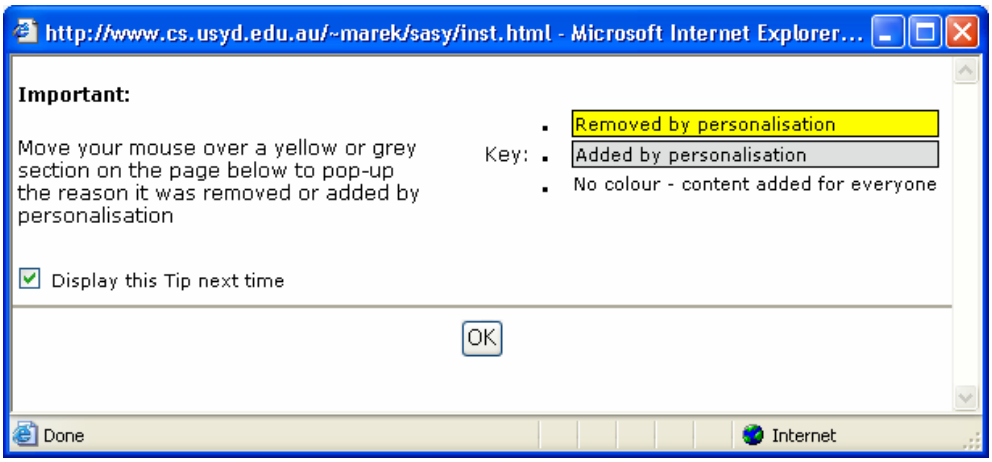


Figure 6.11 – Pop-up added in SASY 3 to provide online help on first access of the Highlight tool.

Iteration 3 – Results

Appendix Table 8.15 (page 250) shows the forms of scrutinisation participants explored in completing the tasks, observation notes are provided in Appendix 8.5.5 (page 266).

Highlight Tool (Scrutinisation A and B)

P10 invoked the Highlight tool and the new window popped-up. P10 first looked at the "Key" on the screen, read it out and said "ok, that makes sense". P10 ignored the instructions on the right of the window but un-checked the checkbox "Display this Tip next time", then clicked "OK" which closed the window. P10 had not read the instructions and did not know about the mouse over tool. P10 could not complete Task 3 until the instructor prompted them to read the instructions. After this the participant easily used the Highlight tool to scrutinise simple and complex forms of personalisation.

Following discussion with the user following the evaluation, P10 stated they did not like the fact they had to use their mouse to reveal explanations. The user suggested all explanations should be visible always. P10 commented "I would prefer to see all the explanations in one go, rather than after a mouse-over. It is hard to remember what the explanation said after it disappears."

Evidence Tool (Scrutinisation C)

P10 first tried looking in the profile, but later found the evidence link without difficulty.

Profile Tool (Scrutinisation D)

As in previous iterations, P10 was confident in viewing and changing their profile.

Qualitative Survey

P10 agreed or strongly agreed with all questions in the survey.

Iteration 4 – Summary

The sole participant in the evaluation of SASY 3 did not notice the just-in-time help screen, we concluded the new screen was not helpful and abandoned this iteration.

Two users (P10 and P7 from Iteration 3), commented they would prefer to use a version of the Highlight tool that displayed all available information when the tool was invoked, rather than requiring the user to "hunt" for the information by using the mouse to pop-up further explanations. Up to this point we had not considered this option, mostly because it would clutter the page. The SASY 3 Highlight tool allowed the user to see all the content available for a page (i.e. as if personalisation was switched off) yet still provided guidance by highlighting the content that was appropriate and inappropriate for the user. We had observed in Evaluation 1 that a small group of users do actually prefer to read pages in this mode under certain conditions. Adding explanation text around every personalisation might annoy these users. However, since there were now two users suggesting this approach, we were keen to see whether it would improve task performance.

6.3.8 Iteration 5 – SASY 4

We made one more variant, SASY 4 (as described in Chapter 3), where we simplified the Highlight tool so that mouse-over interactions were not required. The Highlight tool in SASY 4 displays explanations for added and removed content on the page itself, rather than requiring mouse interaction. The purpose of this iteration was primarily to evaluate the new form of the Highlight tool in SASY 4. We repeated the experiment with three participants P11, P12 and P13.

An example of the Highlight tool in SASY 4 is shown in Figure 6.12. In the figure, the content "SBS – Filipino News" is highlighted in grey (meaning the content had been added by personalisation) and it includes the explanation text "Added because profile has: You are interested in Current Affairs programs". By contrast, in previous SASY versions the user would have to move their mouse over the grey area to pop-up this explanation text. Essentially in previous versions the explanations were initially hidden, popped-up temporarily and became hidden away when the mouse cursor was moved away, whereas in SASY 4 they were always visible.

Iteration 5 – Results

Appendix Table 8.15 (page 250) shows the forms of scrutinisation participants explored in completing the tasks, observation notes are provided in Appendix 8.5.6 (page 267).

Highlight Tool (Scrutinisation A and B)

The results showed a strong improvement over the previous evaluations. All participants quickly and confidently accessed and used the Highlight tool to complete the first part of the worksheet, including seeing what was added/removed to create the personalised view (scrutinisation form A) and accessing explanations of why each adaptation had occurred (scrutinisation form B). In task 7 where personalisation was based on inferred attributes, users did not initially think to use the Highlight tool but rather first tried the profile. This was because the profile was more familiar to them and they believed it the task required them to change one of their initial preferences. P12 and P13 could not work out the answer from the profile and then tried the Highlight tool, which lead them to the correct answer.

Evidence Tool (Scrutinisation C)

P11 and P12 did not realise they could use the Evidence tool to complete Task 10.

Profile Tool (Scrutinisation D)

As in previous iterations, users was confident in viewing and changing their profile.

Qualitative Survey

Participants agreed or strongly agreed with all questions in the survey, except P11 expressed difficulty using the profile because the user did not initially notice the link that accessed the profile in the top menu. This is because the menu disappeared once the page was scrolled down.

Iteration 5 – Summary

The SASY 4 version of the Highlight tool allowed all three participants to easily complete the tasks that required use of the Highlight tool (scrutinisation forms A and B). All participants immediately read the explanation text and correlated this with attributes stored in their profile.

Figure 6.12 – Screenshot of Highlight tool in SASY 4.

6.3.9 Discussion

We first discuss the effectiveness and usability of the scrutinisation tools (Goal 1).

Highlight Tool

Task 2 of the worksheet required participants to find out what adaptive content had been added or removed from a page to create a personalised view (scrutinisation form A). We expected participants would easily find and use the Highlight tool for this purpose. We found, in the Pilot evaluation (Iteration 1), SASY 1 users had difficulty finding the Highlight tool at all. This was due to a confusing user interface, which was corrected in SASY 2. In the following evaluation iterations, all participants except one quickly found the tool and used it to easily complete the task.

Tasks 3, 4, 5, 7, 8 and 10 tested use of the Highlight tool more deeply, to understand what aspects of the user model caused each personalisation (scrutinisation form B). Using SASY 2, although users easily found the Highlight tool, they were not able to access the explanations because they did not notice the mouse-over tool, nor the instructions on how to access it. Once they read the instructions they used the tool effectively, but some complained the tool was tedious to use. Online training (Iteration 3) did improve performance, but users occasionally had difficulty recalling how to access the mouse over explanations, and it is not clear that users would in practice make use of it in a deployed system. Adding pop-up help when the user invoked the Highlight tool (Iteration 4) did not solve the problem, as the user simply ignored that too. In Iteration 5, we created another form of the highlight tool in SASY 4, that eliminated the mouse over interaction altogether. It displayed all scrutinisation information on the page by default when the Highlight tool was invoked, rather than forcing the user to use their mouse to reveal explanations one at a time. Users quickly noticed the explanations and easily completed tasks. Essentially, this is the same issue as we found with the Cell-Tutor system: increasing visibility of the scrutinisation tools made them more accessible to the user.

The SASY 4 form of the Highlight tool was most effective for task performance in this evaluation. However, the mouse over form has the advantage that it annotates the personalised content but does not add clutter to the page. So in some cases, the mouse-over form might be better.

From the qualitative survey, responses for all the forms of the Highlight tool were positive (except for one user). Users believed it was easy to use all forms of the Highlight tool. However, based on their observed ease of use during task completion, it was apparent that the SASY 4 version of the Highlight tool was favoured by users.

Once users understood how to access the Highlight tool, they were easily able to use it to scrutinise simple forms of personalisation (Tasks 2 to 6) that were based only on initial profile preferences. However, there were mixed results for more complex forms (Tasks 7 through 11), that required participants to appreciate personalisation was not only based on their initial preferences, but also attributes that had been inferred about them by SASY. For these tasks, most user (8 users) first

accessed the profile tool to find out more information about inferred attributes, rather than first using the Highlight tool to inspect the personalisation. However, users did comment that these Tasks were "tricky" so participants did not initially know the best way to complete these tasks.

Evidence Tool

Task 10 required participants to use the evidence tool to find out why SASY had inferred a belief about them (scrutinisation form C). Only five out of the thirteen participants (38%) immediately clicked on the Evidence tool links to easily complete the task. All participants eventually found and used the tool to complete the task but required more time to find the tool, and in some cases, prompting from the instructor to examine the interface carefully.

Once users read the explanation text displayed in the Evidence tool, they indicated it was clearly understood. However, since this evaluation only had one task that tested use of the tool, it can not be generalised whether this tool is usable and effective. Clearly it is effective for some users, but for others its purpose is not clear.

Profile Tool

Tasks 6, 7, 8 and 11 required participants to change their profile to control what the personalisation added/removed from their personalised view (scrutinisation form D). All participants confidently completed Tasks 6 and 11and clearly understood that changing the profile would change the personalisation. Tasks 7 and 8 required users to first understand what personalisation had occurred. Interestingly, users completed these tasks in different ways. Some users used the Highlight tool to first work out what had to be changed in the profile (P1, P4, P5, P6, P7, P8, P9, P10, P12, P13). Others first examined the profile (P1, P2, P3, P4, P6, P11, P12, P13). Of these, some took a guess at what to change in their profile without fully understanding the implications (P2, P3, P5, P8, P11), others abandoned the profile when they realised they could use the Highlight tool (P1, P4, P6, P12, P13) to work out why certain adaptations had occurred.

Four users (P4, P6, P7 and P8) had difficulty finding the link to access the profile in the menu, but eventually found it. This was because once users scrolled down a page, the menu at the top of the page was effectively hidden.

It is clear the Profile tool was both effective and users found it easy to use. The tool was not changed during the evaluation. One user commented that a potential improvement might be to allow users to change profile settings from the TV Guide page, rather than having to navigate to the Profile page. This is more evidence that some users prefer to have all information visible and directly accessible when they need it.

User feedback on the notion of scrutability and control

Another goal of this evaluation was to obtain feedback about the notion of scrutability and control in general. It is interesting to point out that based on the survey all users agreed or strongly agreed

- they thought it was useful to be able to inspect and control the personalisation;
- they were in control of the personalisation;
- they understood how the personalisation worked.

P9 commented "Once I understood how the system could be manipulated I was able to gain more control over the programs I wanted to watch." P6 said it is especially useful to be able to inspect and control the personalisation "to change the automatic special interest group allocation". That is, to change the inferences SASY made about their interests.

It is also interesting that some participants expected the system to make inferences about them and update their profile without consent and were comfortable with this. P7 commented the user model should be built unobtrusively, stating SASY "should learn about what programs I like to watch as I use the system. I shouldn't have to tell it". But the user also stated the system should seek approval before making profile changes. P8 and P10 said they had suspected SASY might change their profile when they clicked on programs they were interested in during Task 9. When logging in to SASY for the first time, P6 noticed the section on the profile titled "So far, SASY has inferred the following about you" and stated they understood that SASY would make observations and inferences about their interests later.

On the other hand, some users were upset that SASY made inferences about them without their consent. P5 said they were surprised and upset the system made them a member of the Special Interest Group for religious TV programs. P10 said "System gonna try to learn about me, but it may not have the full info it wants because I might be interested in programs I didn't click on." This user was concerned the system might develop misconceptions about their interests. The user added "I feel that it may make very strong assumptions especially when I am interested in Getaway [a TV program] in general, it made an assumption that I was interested in all programs that had religious content. I may just happened to be interested in what Getaway was offering at that point in time like cathedrals but I may not be interested in Jesus."

Limitations of the Evaluation

(1) Users expressed confusion about the Special Interest Group profile attributes SIG 1 and SIG 2. This was intentionally obscure in order to encourage users to search for the meaning of SIG1 and SIG 2 using the profile and evidence tools. However, it caused confusion for some participants, particularly P7, so it may have affected task performance.

(2) The task oriented nature of the experiment required heavy use of the scrutinisation tools to complete the worksheet. Participants 7 and 8 commented that if using the system in real life, they would not want to have to interact with the tools so often to complete simple tasks such as viewing a TV Guide. They felt personalisation should be setup once and left alone to do its job without requiring attention from the user.

(3) Participants had prior knowledge of the TV programs referred to in the experiment materials. This encouraged them to guess answers to some tasks rather than use scrutinisation tools. Essentially this meant some user's answers did not provide any data about scrutinisation tool usage because they guessed the answer. This was reported as a failed task (marked with a ?) in Table 8.15.

6.4 Conclusions and Summary

We discuss each of the four evaluation goals.

Goal 1 - Can the users work out how to use the scrutinisation tools without training?

Evaluation 1, a field study, demonstrated that users were able to scrutinise an adaptive hypertext. Evaluation 1 did not require users to access training but users were free to do so. Out of the 105 users, 39 read through an online training guide, so the results of Evaluation 1 only provide partial evidence for this evaluation goal. Since the user's main task was to learn about UNIX security, any scrutinisation was an aside activity and a distraction from their core task. In this context, overall scrutinisation was quite high: 81 users (77%) scrutinised in some way, 13 (12%) scrutinised 9 or more times (i.e. more than twice average). The number of users that explored each form of scrutinisation was:

- Highlight Tool – 36 (34%) users inspected which aspects of the hypertext were personalised (i.e. added to and/or removed) and which aspects of the user model caused each personalisation. In the qualitative survey, only 11 users (15%) strongly disagreed or disagreed they knew how to inspect the personalisation to see what items were included or removed from their personalised view.
- Evidence Tool – 45 (43%) users viewed evidence of how user model attributes were set, whether set directly by the user or inferred by the system.
- Profile Tool – 49 (47%) users viewed their profile and 37 (35%) changed aspects of their user model to control the personalisation. In the qualitative survey, only 11 users (15%) strongly disagreed or disagreed they knew that they could change the beliefs on the profile page to change what was included and removed by the personalisation.

Evaluation 2, a qualitative study, explored whether the scrutinisation tools were effective in allowing scrutinise an adaptive hypertext in the following ways, as described in Chapter 1:

A) Understand what aspects of the hypertext were personalised. That is, what parts of the content were added to and/or removed from the user's personalised view of the page.
B) Understand what aspects of the user model caused each personalisation. See the implications of changing their user model, in terms of how it effects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes.
C) View and understand evidence of how user model attributes were set, whether set directly by the user or inferred by the system.
D) View or change aspects of their user model to control the personalisation.

The study found users quickly worked out how to use the Profile tool and understood that changing the profile would change the personalisation. We did not need to change the Profile tool as it was already effective in SASY 1. On the other hand, the study also found users initially had some difficulty using the SASY 1 version of the Highlight tool. However, once they had explored used the interface, users found it easy to use. There were mixed results regarding the Evidence tool. It was effective and easy to use for some, but this was not the case for all participants.

We used user feedback to improve the Evidence tool and Highlight tool from the SASY 1 version to that of SASY 4. The refined tools greatly improved task performance. Therefore, the SASY 4 scrutinisation tools were more effective for untrained users. Using the SASY 4 tools, users were able to scrutinise in each of the ways described above. The usability of the Evidence tool is not conclusive since only one task measured this, and there were mixed results. However, all users eventually understood how to use the tool.

The main limitation of the evaluations is that most of the sample population were quite computer literate. This means we can not generalise that all users would find the scrutinisation tools easy to use.

Goal 2 – Given access to SASY's scrutinisation tools in an authentic adaptive hypertext application, what are usage patterns of the scrutinisation tools?

Evaluation 1 reported how users scrutinised in an authentic adaptive hypertext developed for teaching. Despite users main learning goal, the majority of users (81 users, 77%) devoted time to scrutinisation to inspect and control the personalisation process.

Based on the experiment design, we had expected to see scrutinisation motivated by the desire to understand and correct personalisation when it produces unexpected or irritating results. We expected users would scrutinise the personalisation, evidence and change their profile to:

A) Remove jokes and hints from the material;
B) Remove advanced concepts from the material;
C) Only show a small number of quiz questions included in the material;
D) Reflect the user had some knowledge about a concept though they did not answer every question about that concept correctly in the pre-test.

All these expected forms of scrutinisation were exhibited, but only by a small subset of users. However, responses to the qualitative survey revealed that students did not find the jokes and hints annoying, which perhaps explains why this did not provoke scrutinisation.

In addition, we expected to see other natural forms of scrutinisation though we could not predict what they would be. In fact, there was more scrutinisation due to these natural motivators than the artificial ones we had planted (i.e. scrutinisation due to jokes, hints etc):

- Desire to understand and control the knowledge level that the system modelled about the user in their profile (user model). The study showed users were quite interested in how the system modelled their domain knowledge and exercised considerable control over these and other aspects of their profile and personalisation: 37 users (37%) accessed evidence about their knowledge level, 24 (22%) change the knowledge level in their profile.
- Reflection about the user model, personalisation and use of course map at logical milestones during learning: just before and after a pre/post test, learning module or session. We observed most scrutinisation actions (96%) occurred when the students were either about to start, or had just finished a learning module, session, pre or post test. This is in tune with the notion that although scrutinisation is an aside activity, it is important. Learning UNIX was the main goal but the scrutinisation tools allowed users to see how their profile and the personalisation changed as their knowledge changed, and query the system about its adaptive response in relation to their profile.
- Desire to access all content available rather than just the personalised view. There were 7 users (7%) who changed a system preference in the profile so that the personalisation would annotate, but not remove, content that was inappropriate for them based on their profile. There were also were 7 users (7%) who used the Highlight tool seemingly to get information they were shown previously but had since been removed by the personalisation as their user model changed.

Goal 3 – Obtain user feedback about SASY's scrutinisation tools and on the notion of scrutability and control of personalisation.

The studies also collected user feedback about SASY's scrutinisation tools and on the notion of scrutability and control of personalisation. Responses to the qualitative surveys in both evaluations showed an overall positive attitude to the notion of scrutinisation and control of personalisation. The majority of users agreed it was useful to be able to inspect and control personalisation.

Evaluation 2 captured comments from users about this issue, where we have set up the personalisation so that SASY made general assumptions about users. Some users were comfortable that SASY learned about them through their use of the system. Other users felt the SASY should confirm all assumptions it made about them.

We found participants in both evaluations were quite comfortable with the notion of personalisation driven by a profile they could view and change to control the personalisation. This strengthens the importance of developing scrutinisation tools and building them into the core architecture of adaptive systems.

Goal 4 – Evaluate across different subject matter domains and styles of adaptive hypertext system

Since SASY was designed to provide a shell for creating scrutably adaptive hypertexts, we wanted to evaluate it across varied domains and across different types of adaptive hypertexts.

Evaluation 1 used SASY as a personalised teaching system for UNIX security concepts. Evaluation 2 used it as a personalised TV Guide recommendation system. In Evaluation 1, users were focused on learning and generally scrutinised at logical milestones just before or after completing a learning module or session. Some users used the Highlight tool to get at information that had been removed by personalisation, with some users preferring to see all the information available rather than just their personalised view. In Evaluation 2, since page content was structured as discrete recommendations, the page content did not have to flow from one paragraph to the next as much as it would for a teaching text. Hence users were comfortable for the Highlight tool to display detailed explanations about the scrutinisation on the page itself, and did not feel that this cluttered the page.

Chapter 7 Conclusions

This thesis explored how to provide effective support for user scrutiny of personalisation of an adaptive hypertext. It extends the work on scrutable and open user models, that make personal data available to the user, to involve the user in the personalisation *process* and empowering them to understand and potentially alter it to better suit their needs.

This thesis presented SASY and ATML, a scrutable adaptive hypertext framework and scrutinisation tools that allow the user to, not only access, understand and modify data that is stored about them, but also, to enquire, understand and control how that data is used and processed to produce the personalised response.

The research is important for several reasons:

- to compensate for the lack of predictability, transparency and controllability in adaptive hypertext systems (Jameson, 2003; Maes, Schneiderman, 1997);
- provide tools to support legislation that requires users to not only have access to personal data stored about them, but also the support to understand how that data us used by the system (European Parliament Privacy Directive, 1995; Kobsa, 2002);
- allow users to:
    - better understand their user model by seeing the implications of the data it contains, in terms of the personalisation it controls;
    - experiment with "what-if" scenarios to explore alternate forms of personalisation. For example, consider an adaptive hypertext for teaching. Different users might prefer different levels of information: some might want to learn the core concepts needed to pass the course, others might also want to learn advanced concepts. As another example, students might prefer a details about concepts when first learning them, but only a summary of the key points when revising for an exam;
    - validate aspects of their user model and correct misconceptions and errors the system has about them – this has the potential benefit of improving the accuracy of the user model and, in turn, improving the personalisation;
    - satisfy their curiosity and exploration of how the personalisation works;
- potentially encourage reflection on the personalisation and user model in a learning context. For example, consider an adaptive hypertext for teaching that personalises to suit a particular learning style. The learner might want to see how the presentation of material differs for different learning styles. Exploring and understanding the differences, and controlling which style is presented to them, might help them learn the material and learn about the ways they learn best.

As another example, consider a scrutable adaptive hypertext system that personalises holiday recommendations to the user based on their interests, budget constraints and other personal characteristics. The user sees a personalised holiday recommendation, and might want to get answers to the following questions:

1) What has been adapted to me?
2) Which holidays did it recommend based on the assumption that I like the beach?
3) What other holidays might it have shown me, but it did not because it treated them as irrelevant to me?
4) Why does it think I like the beach?
5) How can I control or alter the adaptation so that it does not show me any more beach holidays?

The scrutinisation tools in the SASY and ATML framework, developed and evaluated in this thesis, allows users to get answers to these types of questions. The first question can be answered by invoking the Highlight tool to view what parts of a personalised page have been adapted. Answers to the second and third questions are provided by the explanations shown by the Highlight tool that explain why each adaptation occurred in terms of the profile attributes that influenced the adaptation. The fourth question is answered by the Evidence tool that explains how each user model attribute was set, whether set directly by the user in the profile or inferred based on rules defined by the ATML author. To answer the last question, personalisation is controlled by changing values in the user profile. The user may also use the Highlight tool to see the impact different values, before actually changing their profile.

The next section discusses the major contributions of this thesis. The section after that describes limitations of this thesis and future work.

7.1 Contributions

This section discusses the four major contributions of this thesis, as presented in Chapter 1 (section 1.2, page 23).

7.1.1 SASY and ATML: a scrutable adaptive hypertext framework

The first contribution is the design and implementation of a framework for scrutable adaptive hypertext that makes it possible for the user to scrutinise the adaptation processes, as discussed in Chapters 3 and 4. The framework is reusable, domain-independent, and applicable to different types of adaptive hypertext system.

Part of this contribution is the Adaptive Tutorial Mark-up Language (ATML), used to create scrutable adaptive content. The other part is SASY, a web application responsible for the presentation and delivery of personalised content to the user. Although there are other adaptive hypertext frameworks, as discussed in Chapter 2, SASY/ATML is the first to include scrutability support in the adaptive hypertext language.

SASY and ATML utilise a standard adaptive hypertext technique, building on it with support for the scrutinisation of the personalisation it provides. The complexity of the personalisation mechanism used by SASY is rather modest, but representative of a typical adaptive hypertext system (Brusilovsky, 2001). SASY stores a model of each user, populated initially by the user's answers to a brief questionnaire. This design was intended to help signal to the user that SASY puts them in control: from the very first contact, the user sees their basic user model.

The user model is later updated and refined by the system through inferences and observations made about the user. The user model is processed by rules embedded in the personalised content to determine whether adaptive content should be shown or removed from the user's personalised view of the page. The rules for populating the user model, updating it based on inferences and the rules for conditionally displaying adaptive content to a user can be defined either as a specific value a user model attribute must have, or more complex expressions developed in python code that declare arbitrary boolean expressions based on a one or more user model attributes.

7.1.2 Scrutinisation Tools

The second contribution is a set of Scrutinisation Tools that were developed and evaluated through this thesis, as described in Chapters 3 and 4:

- **Highlight tool –** allows the user to see what parts of a page have been personalised, understand why the personalisation occurred in terms of the user model values that caused the adaptation, and how to change the user model to steer the personalisation.

- **Evidence tool –** provides explanations about how, why and when user model attributes were set, whether set directly by the user through the Profile tool or inferred by SASY based on observations about the user.

- **Profile tool –** allows the user to view and edit aspects of their profile. Also provides a high level view of the effect of user model values on the personalisation.

These tools satisfy the requirements for a scrutable adaptive hypertext that we defined in Chapter 3. The following subsection describes how each requirement is satisfied.

Requirements for a Scrutable Adaptive Hypertext Revisited

Requirement 1. The scrutinisation tools must allow the user to scrutinise the personalised hypertext in the following ways:

- **A** understand exactly what aspects of the hypertext were personalised;
- **B** see what aspects of the user model caused each personalisation;
- **C** understand and visualise the implications of changing their user model, in terms of how it affects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes;
- **D** view evidence of how user model attributes were set, whether set directly by the user or inferred by the system;
- **E** change aspects of their user model to control the personalisation.

The scrutinisation tools in SASY 4 allow users to scrutinise in each of the ways discussed in requirement 1:

A – The Highlight tool uses colour annotation to show every aspect of the page content that was personalised. Content added by personalisation is annotated with a grey background colour, content that was omitted is annotated in yellow.

B – Every SASY personalised page lists the user model attributes that effected personalisation on that page. In addition, the Highlight tool provides an explanation for each personalised item.

C – The Highlight tool explains the conditions and user mode values that caused each personalisation, describing content that was added and removed from the user's personalised view. The user can thus see how changing the user model will impact what the personalisation adds and removed from a SASY page, without having to change the profile.

D – The Evidence tool provides explanations about how, why and when user model attributes were set.

E – The Profile tool allows user to change their user model to control the personalisation.

Requirement 2. The scrutinisation tools must allow the user to understand and control both simple and more complex forms of personalisation, as defined below.

- *Simple* personalisation:
  - Personalisation based on the value of a single user model attribute.
  - Personalisation based on a user model attribute that was set explicitly by the user, such as a user preference.

- *Complex* personalisation:
    - Personalisation based on the result of complex expressions involving one or more user model attributes.
    - Personalisation based on a user model attribute(s) that have been inferred about the user.
    - Personalisation that changes over time based on a changing and evolving user model.

Tutor 1-3 and Cell-Tutor only supported simple personalisation. SASY 1-4 supports both forms. In SASY 1-4, the user model is updated as inferences are made about the user. The inferences and rules that define when adaptive content is presented or hidden from a user are defined by the author through arbitrary boolean expressions in the ATML documents that define content.

Requirement 3. The user interface should be easy to use without training to support users who may not use the tools often. However, it should also support "power" users who will use the tools often.

We evaluated on the ability of untrained users to complete tasks using the scrutinisation tools in Evaluation 2 (page 178), this is discussed in the next section.

Requirement 4. Text explanations provided by the scrutinisation tools should be generated by the system where possible to reduce the effort in creating content. However, hypertext authors must have the ability to override default explanations if required, for example, to explain complex personalisation rules.

ATML supports this requirement. For the simple form of the *adapt* tag, that personalises content (see Chapter 4), an explanation of the personalisation is generated by the system. For the complex form, where authors define arbitrary boolean rule to adapt content, authors also specify a textual explanation of the rule that end users can comprehend.

Evaluation of the SASY scrutinisation tools

The third contribution of this thesis are two evaluations of SASY, our scrutable adaptive hypertext framework, as reported in Chapter 6. This thesis also reported earlier evaluations of systems developed as part of the exploration of different scrutinisation tools, this was discussed in Chapter 5.

For the evaluations reported in Chapter 6, we had four evaluation goals:

Goal 1 – Can users work out hot to use the scrutinisation tools without training?
Goal 2 – Given access to SASY's scrutinisation tools in an authentic adaptive hypertext application, what are usage patterns of the scrutinisation tools?
Goal 3 – Obtain user feedback about SASY's scrutinisation tools and on the notion of scrutability and control of personalisation.
Goal 4 – Evaluate across different subject matter domains and styles of adaptive hypertext system.

To achieved these goals, we carried out two evaluations, as reported in Chapter 6.

The first evaluation reported a field study, where students used SASY to learn about UNIX Security in an authentic environment. In this study, SASY was used as an adaptive hypertext teaching system. The evaluation provided insights in to the ways users scrutinised their personalised learning materials in an authentic field trial of SASY, where their main goal for using the system was to learn about UNIX Security. The analysis observed usage patterns of the scrutinisation tools and feedback regarding the usability of the tools. The experiment concluded that a large number of users (81 users, 77%), were motivated to scrutinise the personalised hypertext and their user model particularly when there they were surprised by the behaviour of the personalisation. Overall, users scrutinised in each way described in Requirement 1 in the previous section. This indicated users could work out how to use the tools (Goal 1). In terms of Goal 2, our observations shows users had scrutinised in each of the ways we had expected, based on the evaluation design. In addition, there was evidence of natural forms of scrutinisation that we had not motivated, including: a desire to understand and control the domain knowledge that the system modelled about the user in their profile (user model); reflection about the user model, personalisation and course map at logical milestones during learning; desire to access all content available rather than just the personalised view. One of the limitations of this experiment was that participants were third year computer science students, so they had a highly technical background. This means we may not generalise observations to arbitrary populations of learners.

The second evaluation reported a laboratory based experiment where users were users completed tasks around a Personalised TV Guide. Here SASY was utilised as an adaptive recommender system that recommended a personalised television program viewing schedule to users based on their interests. Users were required to complete 7 tasks to explain and control simple personalisations (Tasks 2-6, 11, page 250) and 4 tasks for more complex personalisations (Tasks 7-10, page 250). Although participants using SASY 1-3 had difficulty finding the scrutinisation tools, all three participants in the SASY 4

evaluation easily completed tasks for simple personalisation. The three SASY 4 participants had some initial difficulty completing the complex personalisation tasks (see Table 8.15, page 250), but after carefully exploring the SASY 4 interface and scrutinisation tools worked out how to use the tools to complete the tasks. Notably, four out of the ten participants of the SASY 1-3 evaluation had no difficulty completing the complex tasks. This indicates SASY 4 users could work out how to use the scrutinisation tools without training for simple personalisation, but might need more time to become familiar with the tools to efficiently scrutinise more complex forms of personalisation (Goal 1).

Both studies collected user feedback about SASY's scrutinisation tools and on the notion of scrutability and control of personalisation (Goal 3). Responses to the qualitative surveys in both evaluations showed an overall positive attitude to the notion of scrutinisation and control of personalisation. The majority of users agreed it was useful to be able to inspect and control personalisation. This supports the need for tools to allow users to scrutinise a personalised hypertext.

The evaluations reported in Chapters 5 and 6 demonstrate the applicability of scrutinisation to adaptive hypertexts involving different domains, and created for different types of adaptive systems: adaptive information presentation systems, adaptive recommender systems, as required for Goal 4. The following list summarises the evaluations of SASY/ATML and previous systems created as part of this thesis:

- Evaluations of SASY/ATML (and previous systems) as adaptive information presentation systems:
    - Tutor 1-3, Cell-Tutor (Chapter 5) and SASY 1 (Evaluation 2, Chapter 6) were evaluated in a teaching domain, for UNIX shell commands and file system security;
    - SASY 1 has also been used to develop a scrutable, adaptive museum guide (Kay & Niu, 2005);

- Evaluations of SASY/ATML (and previous systems) adaptive recommender systems:
    - SASY 1-4 was used to present a personalised Television program guide to users based on their interests (see Evaluation 1, Chapter 6).

7.2 Guidelines for tools to support scrutinisation of an adaptive hypertext

The fourth major contribution of this thesis are a set of guidelines for the creation of scrutinisation tools. This thesis explored several user interfaces to allow the user to see the implications of their user model in terms of the adaptations that it effectively controls. This section presents guidelines for the creation of effective and usable scrutinisation tools. These guidelines are based on the creation and evaluation of the different forms of support for scrutinisation, as discussed in chapters 5 and 6.

We first revisit the key challenges for creations tools to support the process scrutiny of an adaptive hypertext. We then summarise the different scrutinisation tools that were explored by this thesis, and conclude with the key guidelines.

7.2.1 Key Challenges Revisited

In Chapter 3 (section 3.3, page 77), we outlined the key challenges for developing scrutinisation tools. We summarise the challenges here and discuss how we addressed them in the next section.

Walk-up-and-use Interface

We would expect users might scrutinise irregularly, based on our previous work (Czarkowski & Kay 2000) as scrutinisation is not user's primary reason for using the adaptive hypertext system. User might go for long periods without scrutinising. However, when there is a suitable trigger, such as unexpected behaviour of the personalisation, the user should be able to work out how to delve into the adaptation, without having been trained to use the scrutinisation tools.

Supporting both Novice and Advanced users

Some users may scrutinise often or want to see what personalisation removed from their view on every page. This type of user needs to be supported as does the casual user.

Visibility

There must be a balance between the obtrusiveness, or visibility, of the scrutinisation tools, and the user's ability to find the tools. One might expect if the tools will not be used often they should be hidden away. On the other hand, users must be able to find the tools when needed.

Explaining personalisation to users

The personalisation must be explained to the user in a way they can readily comprehend. In addition, they must be comfortable with several key concepts surrounding personalisation:

- the system stores data about them;
- this data is based on information they volunteered and it is also based on inferences and system observations of their use of the adaptive hypertext and environment;
- this data is used to decide what content is presented to them;

- additionally, for this thesis, users must understand they are able to scrutinisation the personalisation and change the data stored about them to steer the personalisation.

7.2.2 Summary of scrutinisation tools explored

This thesis began exploring ways to allow users to scrutinise personalisation with Tutor 1 (Czarkowski & Kay, 2000). It employed a very simple approach where the user model and personalisation was based entirely on a attributes provided by the user through a short questionnaire: yet this made a logical starting point for exploring scrutinisation. In its evaluation, Tutor 1 was used as a teaching system. Since the user's main focus was to learn, the scrutinisation tools were designed to be out of view on the page, where they would not distract the user: the link *How was this page adapted to you?* Was included at the bottom of every page. Clicking this added a section to the bottom of the page that explained each personalisation through a bullet point list of included and removed items. The visibility of the scrutinisation tools was low and there was minimal textual explanations provided about the personalisation that occurred. In this evaluation, users did not scrutinise often: 29% examined the adaptivity and 27% of these users went further, checking the actual parts of the hypertext page affected by the adaptation.

We made small enhancements to Tutor 1, including the ability to scrutinise content removed by personalisation, and called this Tutor 2. A qualitative evaluation of Tutor 2 found:

- one of the key difficulties was that users were not able to easily find the tools when required;
- surprisingly, users could not scrutinise and were overwhelmed by the explanations provided by the system, even in this simple case;
- users did not appreciate the purpose of the scrutinisation tools.

It was clear at this stage user that interface issues had to be considered carefully as they impacted the users' ability to scrutinise.

We developed a quite different form of the scrutinisation tools in Tutor 3: it provided a summary of the personalisation as well as details of each adaptation; it included explanatory text for each adaptation rather than a bullet point list. The qualitative evaluation of the Tutor 3 scrutinisation tools found users:

- were comfortable with simple user models (i.e. the user model was only updated explicitly by the user and not by the system) used to capture information to drive personalisation;
- understood that they had input to the personalisation process, but needed convincing that they could control it as well;
- still had some difficulty finding the scrutinisation tools when required.

At this stage we had made the explanations of the personalisation and the scrutinisation tools more comprehensible by using complete sentences to describe these aspects rather than coded bullet points. However, user's had difficulty finding and using the tools in their initial use.

The next system (carried out by Serena Potts and Judy Kay) used a more radical approach, where the user model was displayed on every page to the user. The user model was an always present reminder that personalisation was being performed in the background. This approach increased user awareness of the user model and how it was used by increasing the visibility of the profile scrutinisation tool.

We extrapolated this idea of a highly visible scrutability interface and explored more complex personalisation in SASY 1, where the user model was updated by the system based on observations and inferences about the user. This added another level of complexity to the personalisation and in turn increased the difficulty in developing scrutinisation tools to support this. SASY 1was developed to accommodate explore the role of scrutinisation with personalisation that is representative of real-world adaptive systems. We evaluated SASY 1 through a large scale field trial and a small scale qualitative evaluation. For the field trial, we had set up the content and default personalisation in such a way that we expected users would be motivated to scrutinise and change the personalisation to better suit their needs. In fact, the majority of users (N=105, 82%) scrutinised the user model, evidence and personalisation, scrutinising in each of the ways we expected in the evaluation design. This indicated users were able to find the tools. The qualitative evaluation highlighted some interface enhancements that we applied to create SASY 2,3 and 4 (see Chapter 6). The key enhancement in SASY 4 was a new version of the Highlight tool, called the *default* mode, that was more effective for novice users since it showed explanations of all adaptations following a single mouse click.

In summary, by increasing the visibility of the scrutinisation tools and adding textual explanations to describe each adaptation, and to describe how to use the scrutinisation tools, we created a walk up and use interface with SASY 4. The qualitative evaluation of SASY 4 found users could appreciate that:

- adaptation occurred based on their personal attributes stored in their profile;
- that their profile contained information they volunteered about themselves and information that was inferred about through observations made about them by the system;
- that they could change their profile to control the personalisation.

7.2.3 Guidelines for Scrutinisation Tools

This section presents guidelines for the design of effective and usable scrutinisation tools. These guidelines are based on the development and evaluation of the different forms of support for scrutinisation, as mentioned above and discussed in Chapters 5 and 6.

In general, usability best practice guidelines apply to scrutinisation tools as they do to any user interface. The reader may refer to Nielsen (1994b) for a discussion of such guidelines. The key guidelines, as applicable to this thesis, are discussed below.

Predictability

It must be obvious to the user what the scrutinisation tools do. To a large extent, this has to do with how the tools are presented to the user. For example, SASY 1 had the unpredictable hyperlink "Click to highlight removed/included items on this page" (see Figure 5.10, page 139). In contrast, SASY 4 has the hyperlinks "X items removed" and "Y items added", where X and Y are integers that represent the number of personalised content items removed and added to the page by personalisation (see Figure 3.3, page 82). Users readily understood the second form but had difficulty understanding the first form.

Learnability and Memorability

It is expected that most users will not scrutinise often. It might only be when there is a suitable trigger, such as unexpected behaviour of the personalisation, or, as we have learnt in Evaluation 1, when they have just completed a task. To support infrequent users, the user interface should be easy to use without training. The tools must be easy to learn and remember. For example, the mouse-over mode of the Highlight tool employed by SASY 1 was difficult to learn. Users who had read through the training material had forgotten about the tool soon after training.

The evaluation of SASY in Chapter 6 reported that, for novice users, the most effective scrutinisation tools were those that required least instruction and interaction between the user and system. This is consistent with the "paradox of the active user" (Carroll & Rosson, 1987), that users never read manuals but start using the software immediately.

Visibility

Perhaps the most difficult challenge for developing scrutinisation tools was finding the balance of the visibility of the scrutinisation tools. One might expect that since the tools might not be used often, they should be hidden away. On the other hand, users must be able to find the tools when needed. If the tools are hidden, they become too hard to find and users may never find them when required.

The early scrutinisation tools in Tutor 1 were very unobtrusive. The Highlight tool was accessed through a hyperlink at the bottom of each page, as we believed users would be able to find this when needed but we did not want to clutter the page. However, users were not able to find this link. Those who did find it may not have understood its purpose. The Cell-Tutor system experimented with more a more visible user model, which helped to increase user awareness of their user model and personalisation. SASY applied this idea to its other scrutinisation tools, so that they were in constant view on the page. The most visible scrutinisation tools were determined to be most effective for novice users in our evaluation (see Chapter 6 – Evaluation 2). This is particularly important, given that users currently do not expect to be able to scrutinise computer applications (see Chapter 5) but want to have control the system' adaptive behaviour (Bontcheva et al., 2005; Jameson, 2005; Peng et al., 2005; Cheverst et al., 2005; Bohnenberger, 2005).

Although the SASY scrutinisation tools are visible on every personalised page, the user is not forced to view detailed explanations of adaptation. By default, detailed scrutinisation information must be explicitly requested by the user by invoking the scrutinisation tools. Another work (Bohnenberger, 2005), displayed scrutinisation information to user even when they had not requested it in certain situations: in cases where the potential for the user to deviate from the system's recommendation might risk the ability of the user to achieve their overall task, the system displayed scrutinisation information – the rationale being that understanding how the system arrived at its recommendation and understanding the implication of selecting an alternative might help the user make a more informed choice (as discussed in Chapter 2). However, as described in Chapter 3, the user can configure how much scrutinisation information is displayed on each page by default. This ties in with the next guideline.

Flexibility & Configurability

Flexibility provides the user with control. The field study reported in Chapter 6 discussed usage patterns of scrutinisation tools in SASY. It showed users scrutinised in different ways. Some users were very interested in their user model, others in evidence of how user model values were set, while some were most interested in how pages were personalised. In addition, there were several users who wanted to see all the content available for several pages. Effectively, these users wanted to see all the content available for pages rather then just their personalised view. Supporting these users requires flexibility to be built in to the adaptive hypertext system and scrutinisation tools. It also requires careful consideration of how the adaptive hypertext page might appear. For example, even though these users want to see content that personalisation would remove from their view, they might still want some level of guidance or annotation to inform them as to what the content author considered appropriate or inappropriate content, based on their user model. However, too much annotation would mean the page is too cluttered and unreadable. SASY 4 addresses this through the mouse-over mode of the Highlight tool. The tool uses colour to provide annotation but omits detailed text explanations of the personalisation that occurred. Whilst this might not be appropriate for users who want to quickly understand how a page was personalised (as described above), it is better suited to users who want to turn off personalisation for several pages. For example, in the large scale Evaluation 1 (page 146), there were some users (3 users, 3%) who changed their system preference in the profile so that SASY would invoke the Highlight tool on each page by default. The users then worked through the UNIX course in this mode.

In summary, the tools should be flexible and allow users to scrutinise at different levels. For example, some user might only want an overview of the personalisation. Others might want in-depth explanations of all personalisation that occurred. Both should be possible.

The scrutinisation tools should also be configurable to allow users to control the visibility of the scrutinisation tools. Some users might want to know what is personalised on each page, others might

want to ignore the scrutinisation tools altogether. SASY does this by allowing the user to list the aspects of their profile that influenced adaptation, or hide the list altogether (see Chapter 3, page 87).

The notion of configurability of scrutinisation tools is also in line with related research (Bontcheva et al., 2005; Jameson, 2005; Peng et al., 2005; Bohnenberger, 2005), that essentially concludes it is not possible to please all users all of the time. This is because different users have different attitudes and preferences in terms of how much information they want to see regarding the personalisation and their control over it.

7.3 Limitations and Further work

7.3.1 Limitations of SASY and ATML

Complexity of personalisation reduces comprehensibility of scrutinisation

The complexity of the personalisation is directly related to the complexity of the scrutinisation. This limits our approach in a few ways.

Firstly, our approach is limited to small scale user models, those with less than about 20 attributes in the user model. This is because all the user model attributes are displayed on the Profile page and personalised pages list, so a large user model would have the Profile tool appear complex.

A second limitation is the complexity of rules used for personalisation. Although authors may create arbitrary expressions based on user model attributes, complex rules are difficult to explain to users, especially those from a non-technical background. In our approach the author provides text to explain complex personalisation rules: the text is shown to users through the Highlight tool.

Finally, our approach allows personalisation rules to be nested. However, this also complicates the explanations provided by the Highlight tool and at some point multiple nesting levels would be difficult for users to understand. For example, our approach to scrutinisation might not support multiple nested levels of adaptation rules such as is typically used for natural language generation.

Authoring

Our approach increases the burden on the author. The majority of the burden is to provide information used for adaptation. No additional work is required to allow users to scrutinise simple forms of personalisation. However, additional text explanations are required for complex personalisations. It is envisaged that an authoring tool would support this process, but this is beyond the scope of this thesis. On the other hand, the scrutinisation tools actually aid the authoring of adaptive hypertexts by allowing the user to understand the resulting personalisation, an to test and debug the adaptation rules.

End User Population

In our evaluations of SASY and ATML, the majority of participants were highly computer literate. This group has the best chance of understanding and effectively using scrutinisation tools. However, it is not clear how easily less technical users would be able to scrutinise. This is one area for further work.

7.3.2 Further Enhancements to SASY Scrutinisation tools

Tighter integration between understanding personalisation and controlling it

Currently SASY requires users to navigate to the profile page in order to update their user model. An alternate approach is that the user might update their user model directly on the adapted page. For example, Tsandilas & Schraefel (2004) support this. The advantage is a tighter coupling between the scrutinisation and control functions. Such an enhancements was suggested by one user in Evaluation 2.

More Complex Forms of Scrutinisation

Another area of enhancement would be to allow the user to select a single user model attribute, and view the effects of only that attribute on personalisation.

7.3.3 Scrutinising other forms of Adaptive Hypertext

This thesis focused on supporting scrutability for one form of adaptive hypertext: adaptive presentation where personalisation involves inserting/removing content fragments based on logical expressions involving values of user model attributes. In general, the mechanism for generating a personalised hypertext might be more complicated than the rule-based approach used in this thesis.

There are other well known adaptive presentation techniques such as altering content fragments, stretchtext, sorting fragments, dimming fragments. There are also many forms of adaptive navigation support not explored in this thesis (Brusilovsky, 2001). Applying scrutinisation to these forms is one area for further work. One related work has explored the scrutinisation of fisheye adaptation (Tsandilas & Schraefel, 2004) in a recommender style system, as discussed in Chapter 2.

Our approach for scrutinisation could be directly applied to adaptive hypertext based on stretchtext, alternate content fragments, and dimming fragments. However, it is not obvious how to explain and allow the user to control the sort order of content or hyperlinks.

7.3.4 Scrutinising other forms of Adaptive Interaction

Aside from adaptive hypertext, scrutinisation might be explored in other types of adaptive systems. Our approach might apply where a particular item of content was shown to the user based on an a rule or condition that can be explained to the user. However, for other more complex forms of personalisation, such as natural language generation systems, further research is required to find tools to allow user to scrutinise this.

The guidelines for the development of scrutinisation tool do apply to other forms of adaptive systems. For example, users need to be able to easily find and access the scrutinisation tools so the tools need to be visible if users will be able to find them. The adaptive system might also offer several levels of scrutinisation, that the user should be able to configure.

7.3.5 User Model Servers

There has been research in to sharing user models between applications. Essentially, the user model management is contracted out to a third party application, called a user model server. This has several benefits including:

- Reusability of the user model. If the user accesses multiple adaptive systems, there might be potential for these applications to share and maintain the user model, rather than separately model the user's characteristics. This allows for a single, consistent model rather than several different models. This is becoming more relevant with the uptake of mobile technology (Kobsa, 2001). One challenge in this approach is how to recognise an attribute to mean the same thing in several, separate systems. Perhaps scrutinisation tools could be used to indicate to users how the same attribute is used differently by different adaptive systems they interact with. Another application of scrutinisation tools might be to indicate which attributes of their full user model are used by each adaptive system they interact with.
- Access to more forms of scrutinisation. If the user model is independent from the adaptive system, the user might wish to use special purpose tools to visualise and scrutinise their user model. For example, VIUM (Kay & Uther, 2003) is one such tool for visualising large user models. Kay's um toolkit (1999) is another toolkit that allows users to scrutinise their user model.

SASY manages its own user model. Predominantly, this is because this thesis focuses on the scrutability of personalisation, which requires information from the user model combined with information about how the user model was used in a specific context to personalise a hypertext document. However, an extension of this work could abstract the user model to a separate component, such as Personis (Kay et. al., 2002). Equally, SASY could be extended to import or export user models, to make use of a user model server.

7.4 Summary

As part of this thesis, several forms of scrutinisation tools were created and evaluated. Based on evaluation of these tools, this chapter presented the key guidelines to be considered for the design and development of tools to allow users to scrutinised an adaptive hypertext: the tools must be predictable, learnable, highly visible but not distracting, flexible and configurable.

The user of an adaptive hypertext system should be able ask *How was this material personalised to me?* and to be able to see just how their user model affected what they saw in the hypertext document. With an understanding of the personalisation process and the ability to control it, the user is able to steer the personalisation to suit their changing needs. This thesis explored several forms of tools to allow the user to do this and demonstrated that, over very different domains, they can do just this.

Chapter 8 Appendix

8.1 ATML Specification

8.1.1 Profile Page Tags

Course tag
The purpose of the *course* tag is to specify the unique identifier of the topic to which the profile page belongs.

Table 8.1 – Attributes of the Course tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
id	String	1	1	Uniquely identifies the topic internally. For example, "UNIX" could be used for a topic about UNIX.

Initial tag
All user model attributes (*attr* tags) contained by the *initial* tags are displayed to the user as a questionnaire when they first log in to the topic. The questionnaire is used to directly elicit from the user, personal characteristics that are used to build the initial user model. The questionnaire is also shown when the profile is accessed later to allows users to change their responses.

The user model attributes (*attr* tags) contained by the *initial* tag are displayed under the heading "Your Initial Preferences" on the profile page. This heading is automatically generated by SASY.

The *initial* tag contains one or more *attr* tags. An example of an initial profile page is shown in Figure 3.2.

Inferred tag
All user model attributes (*attr* tags) contained by the inferred tags are not displayed to the user in the initial profile questionnaire. Instead, these attributes are inferred about the user or have values defaulted on initial profile creation. The inferred attributes are displayed to the user the second time the profile page is accessed, or on subsequent login. The user is then able to view evidence about how the values were set and change them, should they wish to do so. The user model attributes (*attr* tags) contained by the *inferred* tag are displayed under the heading "So far, SASY has inferred the following about you". This heading is automatically generated by SASY.

An example of how inferred attributes are displayed is shown in Figure 3.9.

The *inferred* tag contains one or more *attr* tags, and one or more *heading* tags.

Attr tag
The *attr* tag is used to define a user model attribute. It defines the metadata for the attribute, including the data type, how the value is elicited, the range of values it may take and information that is provided to the user to allow them to scrutinise the meaning of the attribute on the profile page and on content pages.

Table 8.2 – Attributes of the Course tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
id	String	1	1	Uniquely identifies the user model attribute. For example, "budget". This is only used internally and is not shown to the user.

Attributes	Data Type	Min Occurs	Max Occurs	Description
Acquisition	String	1	1	Defines the data type and how the user model attribute is populated: • "direct" – attribute is a fixed value attribute that is shown to user on initial profile creation. • "inferred" – attribute is a fixed value attribute that is not shown is not show to user on initial profile creation. It is inferred by the system on profile creation (see *init* tag) or set later (see *updateum* tag). • "quiz_score" – attribute is used to hold the value of a quiz score. The score starts at zero and is incremented if a quiz question is answered correctly – see *question* tag. • "prob" – attribute is used to hold the value of a probability integer.
Private	String	0	1	If specified as "true" , the attribute is not shown to the user and the user may not change its value. This can us used for system variables.

Heading tag
The *heading* tag adds a row to the profile page that contains a given text. For example, the profile page shown in Figure 3.9 contains the headings "Additional holiday preferences" and "Your System Preferences".

Table 8.3 – Attributes of the Text tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
text	String	1	1	The text for the heading to be inserted.

Init tag
The *init* tag is used to define a rule that specifies how an inferred user model attribute should be defaulted or updated when the profile page is saved by the user in SASY.

Table 8.4 – Attributes of the Init tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
id	String	1	1	The identifier of the user model attribute that is to be set by the rule.
Cond	String	1	1	The condition that, if true, will set the user model attribute identified by *id* to the specified *value*. The *cond* attribute can be either: • "default" – means the rule is always fired. • A complex expression, as described above. For example, "self.budget=='low'"
value	String	1	1	The value the user model attribute will be set to.
Evidence	String	1	1	A text description that will be displayed by the SASY Evidence tool to describe how the user model attribute was set.

8.1.2 Content Page Tags

Page tag
The *page* tag encloses all the content for a SASY page. The *umKeys* attribute is used to specify a list of user model attributes that will be used by the *adapt* tags on this page. This list is used by SASY to retrieve and display the list of beliefs about the user that effected personalisation on the current page. The list is shown on the right hand side of the page. For example, Figure 3.3 (pg. 82) shows the beliefs "You are a Couple" and "You have a low budget".

Table 8.5 – Attributes of the Page tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
umKeys	String	1	1	Comma separated list of user model attribute identifiers that will be used for personalisation on this page.

Adapt tag

Table 8.6 – Attributes of the Adapt tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
cond	String	1	1	If type is not "complex", *cond* references the identifier of a user model attribute and *value*

Attributes	Data Type	Min Occurs	Max Occurs	Description
				specifies the value that attribute must have for the enclosed content to be added to the user's personalised view of the page. Otherwise, specifies a complex expression that if evaluates to true will cause enclosed content to be added to the user's personalised view of the page.
value	String	0	1	See above.
Type	String	0	1	If set to "complex", specifies the complex form of the adapt tag is used.
Inc_explain	String	0	1	Used if type is "complex", specifies text shown by the Highlight tool explaining why the enclosed was included by personalisation.
Exc_explain	String	0	1	Used if type is "complex", specifies text shown by the Highlight tool explaining why the enclosed was removed by personalisation.

Link tag

Table 8.7 – Attributes of the Link tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
text	String	1	1	The label describing the anchor.
Src	String	1	1	The URL (location) of the referenced ATML page.
pre	String	0	1	A comma separated list of pre-requisite pages (i.e. src attributes) that should be visited by the user before the current anchor.

Pagelink tag
The *pagelink* tag is used to provide navigation buttons to help users navigate between pages. A *pagelink* tag may include a left arrow button hyperlink to the previous page (*prev* tag) and a right arrow button hyperlink to the next page (*next* tag). An example is shown in Figure 4.5 (pg. 108).

Table 8.8 – Attributes of the Prev and Next tags.

Attributes	Data Type	Min Occurs	Max Occurs	Description
title	String	0	1	The label describing the anchor that is displayed next to the arrow button.
Src	String	0	1	The URL (location) of the referenced ATML page.
auto	String	0	1	Applies to prev tag only. If the attribute auto=”true” is specified, the src and title attributes for the previous page are generated by SASY based on the user’s navigation history. This causes the previous page hyperlink to point to the last page accessed by the user.

Updateum tag
This *updateum* tag is used to conditionally update the value of a user model attribute. It is similar to the *init* tag that is used to set initial user model attributes.

Table 8.9 – Attributes of the Updateum tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
attr	String	1	1	The identifier of the user model attribute to be updated.
Value	String	0	1	The value that the user model attribute will be set to. Either a *value* or *inc* attribute must be specified/
inc	Integer	0	1	Only used to update prob type attributes. An amount that is to be added to the current value for the attribute.
Strength	String	1	1	Indicates the certainty of the update: "direct" – full certainty. "inferred" – an inference, not fully certain.
Evidence	String	1	1	Text that is showed to the user through the Evidence tool to explain why the user model was updated.
AddEvidence	String	1	1	Specifies when to add evidence: "ifChanged" – only add evidence if the user model attribute value was changed. "always" – always add evidence.
IfCond	String	0	1	If specified, this condition must evaluate to true in order for the user model change to be applied. This is a complex expression that follows the same rules as defined for the *init* and *adapt* tags.

Getumval tag
This *getumval* tag is used to retrieve the value of a user model attribute so that is can be rendered as part of the content page.

Table 8.10 – Attributes of the Getumval tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
attr	String	1	1	The identifier of the user model attribute to be retrieved.

Question tag

Table 8.11 – Attributes of the Question tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
type	String	1	1	Constant "choice" is the only type currently supported.
Id	String	1	1	A unique identifier for the quiz question.

Table 8.12 – Attributes of the Choices tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
correct	String	1	1	Specifies which answer is the correct answer. This must match the option attribute of the correct *choice* tag.
Scoreattr	String	1	1	The name of the user model quiz_score attribute that is to be incremented if the user correctly answers the question.
Attempts	String	0	1	The number of attempts the user is allowed to answer the question. By default, the user has an unlimited number of attempts that are not penalised.
Concept	String	0	1	The name of the user model prob(ability) attribute that is to be incremented if the user correctly answers the question.

Table 8.13 – Attributes of the Choice tag.

Attributes	Data Type	Min Occurs	Max Occurs	Description
option	String	1	1	Identifier for the answer option.
Umupdate	String	0	1	The URL of the ATML page that is processed in the background when the user selects this answer. The called ATML page may contain rules (i.e. updateum tags) that modify the state of the user model when the user select this answer.
Prob	String	0	1	Specifies a positive or negative integer amount that is added to the probability of the user model attribute specified in the *concept* attribute of the *choices* tag. That is, when the user selects this answer the probability they

Attributes	Data Type	Min Occurs	Max Occurs	Description
				know a concept can be increased or decreased.

8.2 Tutor 2 Evaluation Participant Observation Notes

(Researcher' s comments in italics, student comments in normal font. An interviewer was hired to observe and interview participants as they used the system. Participants were asked to think-aloud as they used the system. Their actions were observed and comments were recorded)

Participant: S (Not Trained)
Date: 21/01/2002
Start time: 10am
End time: 11am

S was by far the most confident of the users and picked up how to use the program very quickly. She was not instructed as to how to use the program.

After approximately 10 minutes of using the program (and making comments for improvement), I prompted her to check her user model. From the UNIX components section, she interpreted "All" to mean all the stuff on shell or kernel rather than that she had selected to "work through all the topics in detail" when she set up her profile. However, she corrected this misconception when she realised that she could click on the link to change the values in her profile. She also correctly understood the function of highlighting but commented

S – "there should be different colours for the different sections of highlighting" so that you can highlight multiple things and see their relevancy to each other. Suggested different colours with a key

- *I had the thought that if this was to be implemented, the appropriate colour could be shown over the link, eliminating the need for a key*

S – suggested that "How was this page adapted to you?" should be changed to something like "Your Preferences" because people would think that the former would indicate how the tutor was written etc (a bit like a "About this program" readme) and not look at it.

S started navigating from the course map (with prompting from me) after changing the value of what she was interested in her profile from "shell" to "kernel" and not being taken to the start of the "kernel" section. She liked the course map interface much more and suggested that the whole tutor should be navigated from it. However, I noticed that she did not really consider the colour coding of the course map.

There was then a fire alarm from about 10:25 until 10:40 where we had to evacuate the building

S – "There is too much information to read through on how to use the interface. I wouldn't read through all that – I'd just start using the program"

When asked about her overall impressions were, S said that she would probably not access how the page was adapted to herself. She said that the adaptation interface was confusing, even though I observed her using it seemingly without any trouble. I believe that she understood how to use the tutor but did not rate the adaptations as an important feature. She suggested some interface changes, including compulsory viewing of the key to the user model, saying that inexperienced computer users would not be able to navigate it without help. She suggested that reordering the pages to include this information at the start would be helpful.

Participant: AC (Not Trained)
Date: 21/01/2002
Start time: 11:15am
End time: 11:50am

C was very confused by the adaptations initially, however after a considerable amount of prompting and asking questions, she came to realise what the functions of the adaptations were. However, I suspect that if she were not prompted to think about their functions, she would not have thought about

them and/or used them at all. AC was given no instruction as to how to use the tutor but was asked to access the adaptations. However, she did not do so regularly unless prompted.

I asked her if she understood the function of the links in "page content included". She did not –

AC – "Aren't they supposed to take you to the individual topics"

Even after several tries at clicking on the links and being redirected to the "Your Profile" page, she still did not seem to understand the function of these links

11:30
When I asked her about the function of highlighting, she responded that it was:

AC – "Whatever we're reading at the moment"

Which I understood to mean what subject we were reading at the time and nothing to do with her adaptation. I reminded her at this point to please let me know if her opinions changed at any time.

She also said at this point that she was very confused. She still believed that the adaptations served the function of the course map – navigation and as a table of contents.

11:33

I suggested that she change from "skim through all topics" to "work through all topics in detail" in her preferences to see if this would help her understand how the highlighting works (as on "skim" only the navigation seemed to be highlighted – which would explain in part the comments she made above at 11:30) – I hoped that she would realise that content could also be highlighted.

This seemed more successful but she was still confused, saying that content highlighting showed

AC – "The most important parts" [of the text]

But still did not realise that it was adapted to what she had specified in her preferences.

She did not intuitively use the unhighlight link, but rather thought that the number of things highlighted increased as you continued through the tutor

However, at 11:40, she seemed to realise correctly what the function of highlighting was!

When asked to access "Show all text" and to let me know what she thought, she replied:
AC – "It lets you see all the navigation options"
And for "Adapt…"
AC – "Shows you the present option"
She seemed to think that it was restricted to showing the navigation options rather than including any extra text/questions etc

I noticed that she did not access "See help" even though she said that she was confused

AC – "Why is there sometimes only one opportunity to highlight and sometimes a lot?"
After some discussion she seemed to think that as you continued through the program and built up more information, more highlighting options were made available. I thought that she had understood the function of highlighting a few minutes previously (at 11:40)

However, a few minutes later she understood that it was in response to the *"required value"* and her own choice

After asking her about the functions of the following, she replied:

Page content included: *"Where you can get the text information"* which I took to mean a link to the appropriate information rather than a link to change your preferences

Page Navigation Excluded: *"Go to the options and say what it's all about"* which I understood as the ability to change your preferences
Page navigation included: *"Navigation buttons"*

She stopped using the program at approximately 11:46
A – "Would you access the adaptations yourself when you use the tutor?"
AC – "Now that I know what to do, I would use it but I wouldn't have used it before – I just would have clicked on the navigation buttons. I would have been curious to what it meant though because it seems more personal"
A – "What changes would you make?"
AC – "Combine where you are now and where you want to go" which I took to mean combining page navigation and content – however, as I don't know that she entirely understood the difference between them I don't believe that this is exactly what she really meant.

Participant: P (Trained)
Date: 21/01/2002
Start time: 11:59am
End time: 12:35pm

I explained the program and how the adaptations work until approximately 12:06. Showed her how highlighting etc works. Asked her to please use the adaptations as much as possible and let me know what she thought at each step.

She seemed to understand what was going on
12:13
P – "It doesn't really need the highlighting function, as I can figure out for myself what the appropriate examples are"

She liked the step by step approach and kept commenting on this during her use of the program. She continued to use highlighting/unhighlighting and seemed to have a good grasp of this.
When asked, she said that she was interested to know what was adapted to her

~12:20
P – "The amount of information is overwhelming both when it is adapted to me and is general. Would be much better if it was broken up into chunks and dealt with in smaller steps"

12:23 – when asked, she said that it didn't occur to her to change the values in the adaptations to change what was displayed to her but that she would consider doing so now that it had been suggested

12:26
P – "I probably wouldn't keep clicking this link [How was this page adapted to you?] if you didn't ask me to – they all seem the same anyway"
But commented that if she was interested in a particular goal (like learning about files) than just learning UNIX, she would probably find the adaptations more useful

She did not attempt to change any values during the walkthrough even though she seemed to understand that the output depended on what she had specified in those options

She stopped using the program around 12:30

P – "I probably wouldn't use the adaptations if working on my own"
When asked, she didn't think that she would need to change them after she defined them at the start but admitted that if she did need to change them that they would be useful.
She said that highlighting was most useful in relation to the questions as the highlighted bit was often helpful in answering them.

P – "I like that the adaptations are not intrusive – you don't have to see them if you don't want to"

Participant: P2 (Trained)
Date: 21/01/2002
Start time – 12:59pm

End time – 1:32 pm

I explained the program and adaptations until about 1:04pm. I asked him to use the program as normal including clicking on adaptations as he saw fit – explained that we were interested in how users responded to the adaptations.

1:14pm
A – I've noticed that you have been using the program for about 10 minutes and you haven't viewed your adaptations yet. Is there a reason for this?
P – " I just want to skim through. I want to see what's available before I see what's missing"

P – commented that the descriptions in the program are too length and was too much to digest at once

I have noticed that all the students commented in some way about the length of the descriptions/examples/etc. However, it was not intuitive for them to downsize the number of examples or choose to only have an abstract definition of UNIX examples

1:20pm
I noticed that after he accessed the quiz page, P then went back a few pages and then changed his profile values through viewing the adaptation. I asked him why.

P – "I want to see the information that's displayed that's different to what I've seen" – he wanted to see the different examples given and different questions to get a different point of view and see if there was any different information that he had missed out on.

I then asked him why he hadn't just chosen to display all text in the adaptations rather than changing his whole profile.

P – "I wanted to change it to what I want rather than see all that's available"

1:25
P – complained that the text was too dense and he could not find the relevant information
A - *I asked him if he considered using the highlight function to find the relevant information*
P – said that he "definitely" had considered this prospect

However he did not do so.

At this point he stopped using the program.

After finishing use of the program, I asked him to comment on the usefulness of the adaptations. When asked about highlighting, he commented
P – the highlighting would be useful in displaying relevant bits of information on the screen even though he had not done so himself
When asked about exclusions, he commented
P- did not care about knowing what wasn't there and wouldn't change his options based upon what was listed as excluded but relevant to the page
P – Would not add or delete anything from the adaptations. Suggested that the explanations and examples were long winded in the actual program

When asked whether he would use the adaptations when using the tutor in his own private study
P – "I would most likely skip through the adaptations but there are some exceptions. For example, if I found a topic interesting, I might turn on the historical aspects option to learn more about it. If I wanted more examples I would use it to change to multiple examples and same for questions"

8.3 Tutor 3 Evaluation Worksheet

For the purpose of this task, imagine you are Fred.

About FRED…

Fred is new to UNIX and wants to learn common UNIX commands. He does not know anything about UNIX. He does use a computer at home which runs Microsoft Windows but he does not know anything about MS DOS.

Fred only has a one hour to learn some UNIX commands so he does not have the time to learn about any other UNIX topics. He is not at all interested in historical aspects of UNIX at all.

Fred usually likes to learn by seeing lots of examples, but also likes abstract definitions. Fred hates being asked questions to prove his knowledge – he finds them irritating.

PART 1

- Open up your Netscape browser and navigate to http://www.cs.usyd.edu.au/~marek/Tutor3

2. Register and log in with the username assigned by your supervisor.

3. Select the course **An Introduction to UNIX**. Read and accept the terms and conditions.

- You should now be on a page titled **Your profile**. Put yourself in Fred's shoes. Using the description about Fred above, fill out your profile as if you were Fred and were in his situation. **(Before you click OK please answer the following questions)**

- Will Tutor use Fred's profile settings? If so, what will it use Fred's profile settings for?

- Where does Tutor get information about Fred to use to perform adaptation of the content to suit him?

- Will Fred be able to influence or control the way Tutor adapts content to him? If so how?

Click **OK**. The page displayed is the **Teacher's Instructions** page. Look over this page and then move to the questions in Part 2.

PART 2

The following questions are based on the **Teacher's Instructions** page. Take some time to examine the page and then answer the following questions.

- What would you do in the system to find out whether Tutor adapted any material on the **Teacher's Instructions** page to you?

9. Did your answer to the profile question "What is your main objective?" have any effect on the contents of the **Teacher's Instructions** page?

(circle appropriate) YES / NO

How do you know this?

10. Was any content specifically excluded because of your answer to the profile question "**What is your main objective**"?

(circle appropriate) YES / NO

If so, what was the content?

11. Was any content specifically included because of your answer to the profile question "**What is your main objective**"?

(circle appropriate) YES / NO

If so, what was the content?

12. Now consider what the first sentence on the page would read had you answered the profile question "**What is your main objective**" with "**Revise the material**".

Without changing your answer just yet, write out how you think the sentence would read…

By the end of this course you will have …

13. Explain what actions you would have to perform in the system to change your answer to the profile question "**What is your main objective**"?

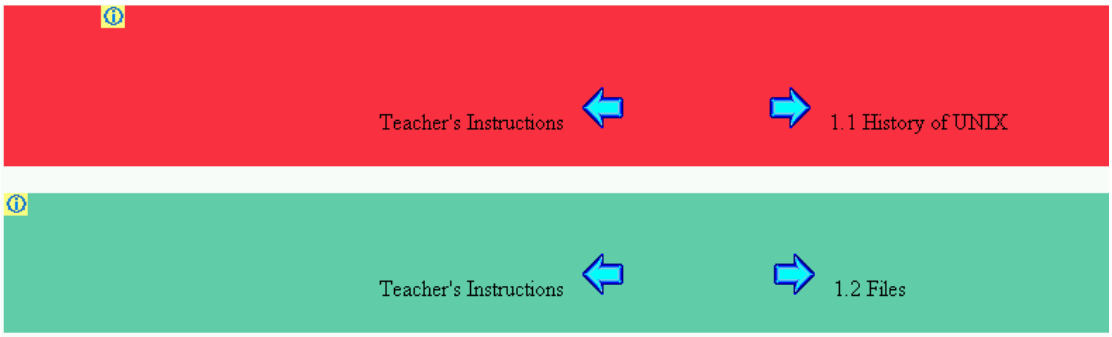
14. Change your answer for the profile question "**What is your main objective**" to "**Revise the material**".

15. Now that you have changed your profile, does the **Teacher's Instructions** page appear as you predicted it would in Question 12 above? If not, can you explain why?

PART 3

Navigate to the page "**1.0 Introduction to UNIX**".

On page "**1.0 Introduction to UNIX**", examine the explanation of what content was included/excluded based on the profile question "**Would you like to see historical background and details of the reasons for the design of aspects of Unix**". The explanation which highlights what was included/excluded should look something like the following:



16. Why are two sets of navigation arrows shown here in the explanation when on the main lesson page "**1.0 Introduction to UNIX**" you were only shown one set?

17. Is the first navigation arrow set included or excluded?

18. Why was the first navigation arrow set included or excluded?

19. Is the second navigation arrow set included or excluded?

20. Why was the second navigation arrow set included or excluded?

21. a) Describe the actions you would perform on the system to make the excluded navigation arrow set become included.

b) If you were to make this change, would both navigation arrow sets then appear on the page "**1.0 Introduction to UNIX**"? Why?

PART 4

Thank you for your participation.

This section is to allow you to give some feedback. Your comments are greatly appreciated.

Did you understand how the system was adapting the course material to you?	Circle one of the following: 1 2 3 4 5 6 7 (No) (Yes)
Describe how you think the system adapts the material to you?	Answer
Did you feel in control of the adaptation? That is, did you feel you had power to change the adaptation to suit your needs.	Circle one of the following: 1 2 3 4 5 6 7 (No) (Yes)
Tutor explains how it adapted a page to you by highlighting sections of content that are included/excluded. These explanations are called **adaptation explanations.** Did you find adaptation explanations easy to understand?	Circle one of the following: 1 2 3 4 5 6 7 (No) (Yes)
Do you think these explanations could be improved? If so how?	Answer:
If there is one thing you would change about Tutor's adaptation explanations, what would it be?	Answer:
What was the best thing about using Tutor?	Answer:
Do you have any comments about this experiment (for example, too long, too short, were any instructions unclear?)	Answer:

8.4 Evaluation 1 – Field Study

8.4.1 Pre-Test Questionnaire

Question 1) What is the purpose of a UNIX shell?

- Allows you to run programs that perform tasks through the operating system.

(B) The shell is another name for the UNIX kernel

I A piece of hardware of a UNIX System

(D) Don't know

Question 2) What is the UNIX shell command to list the current directory contents and also show security access rights set on each object?
- dir

(B) ls
I ls –l
(D) Don't know

Question 3) Which of the following is a standard UNIX ksh environment variable?
- shell

(B) SHELL
I Shell
(D) Don't know

Question 4) Which of the following shell commands correctly adds a directory called dir to your current PATH?
- PATH=$PATH+dir

(B) PATH=$PATH:dir
I PATH=$PATH,dir
(D) Don't know

Question 5) Assume you are in the /usr/bin/programs/ directory. Which of the following UNIX commands will NOT change the working directory to the top level root directory in UNIX?
- cd /

(B) cd root
I cd ../../..
(D) Don't know

Question 6) Assume you are in the /usr/bin/programs/ directory. What is the absolute path to the root directory?
- /

(B) ../../..
I ../..
(D) Don't know

Question 7) Assume you are in the /usr/bin/programs/ directory. What is the relative path to the root directory?
- /

(B) ../../..
I ../..
(D) Don't know

Question 8) A directory called tmp has the following security details as returned by a UNIX list command:
drwxr-x—x 2 jen simpson 512 Nov 02 06:19 tmp
Who is the owner of the directory?
- jen

(B) simpson
I jen simpson
(D) Don't know

Question 9) What is the UNIX shell command that sets the permission on a file so other users (i.e. not the file owner or those in the owner's group) can no longer write to the file called f?
- chmod 125 f

(B) chmod a-w f
I attrib w f
(D) Don't know

Question 10) What is the UNIX shell command that sets permission on a file called f to be rwxr-xr-x?

- chmod 577 f

(B) chmod 755 f
I chmod 777 f
(D) Don't know

8.4.2 Post-Test Questionnaire

Question 1) Which of the following is not a UNIX shell?

- bash

(B) csh
I kash
(D) Don't know

Question 2) Which of the following is a standard UNIX ksh environment variable?

- path

(B) Path
I PATH
(D) Don't know

Question 3) Which of the following shell commands displays all your current environment variables?

- echo $SHELL

(B) echo $env
I env
(D) Don't know

Question 4) What is the UNIX shell command to display the name of the current working directory?

- path

(B) pwd
I dir
(D) Don't know

Question 5) Which of the following is not performed or managed by the file system?

- file/directory access and privilege management.

(B) reading file data
I ensuring your security file security levels are set as high as possible
(D) Don't know

Question 6) What is the command to change into the bin directory which is a sub-directory of the current directory?

- cd bin

(B) cd /bin
I cd ../bin
(D) Don't know

Question 7) Assume you are in the /usr/bin/programs/ directory. What is the relative path to the /home/fred directory?

- /home/fred

(B) ../../../home/fred
I ../../home/fred
(D) Don't know

Question 8) What are the groups of users for which file permissions can be specified

- file creator, user's group, everyone

(B) file owner, user's group, everyone
I as many groups as you want
(D) Don't know

Question 9) What is the UNIX shell command that sets permission on a file so that every user can write to the file called f?

- chmod 323 f

(B) chmod u+w f

I attrib w f
(D) Don't know

Question 10) What is the UNIX shell command that sets the permission on a file so only 'other' users (i.e. not the file owner or those in the owner's group) can no longer write to the file called f?
- chmod 222 f

(B) chmod o-w f
I attrib w f
(D) Don't know

The following questions are not assessed, we just want your opinion to help improve the system. Please be constructive and comment about SASY the system, not the course or the subject matter.

Question 11) Recall Henrik's hints, the comments in red font which appeared throughout the course material. Which statement describes how you felt about Henrik's hints?
(A)They were annoying. I wanted to get rid of them but eligi know how.
(B)They were annoying, I knew how to get rid of them but didn't feel compelled to do so.
II noticed them but they were NOT annoying.
(D)I did not notice Henrik's hints

Question 12) Recall the jokes that appeared in blue font throughout the course material. Which statement describes how you felt about them?
(A)They were annoying. I wanted to get rid of them but eligi know how.
(B)They were annoying, I knew how to get rid of them but didn't feel compelled to do so.
II noticed them but they were NOT annoying.
(D)I did not notice the jokes.

Now consider whether you agree with the following statements:

Question 13) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.
(A)Strongly Agree
(B)Agree
INeutral
(D)Disagree
(E)Strongly Disagree

Question 14) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.
(A)Strongly Agree
(B)Agree
INeutral
(D)Disagree
(E)Strongly Disagree

Question 15) I feel it is useful to be able to inspect and control the personalisation.
(A)Strongly Agree
(B)Agree
INeutral
(D)Disagree
(E)Strongly Disagree

Question 16) If you did NOT inspect the personalisation, what was the main reason.
(A)Did not feel compelled to
(B)I trusted the system's default personalisation
Iother

Question 17) If you DID inspect the personalisation, what was the main reason.
(A)Curiosity
(B)I didn't like how a page was personalised to me

IA belief about me was wrong
(D)other

Question 18) Do you have any other comments?

8.4.3 Table of scrutinisation of each user

User ID	(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(D) Viewed Evidence from Regular Page	(E) Viewed Evidence from Profile Page	(F) Used Highlight Tool	(G) Changed Profile attributes
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0
67	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
25	1	1	0	0	0	0	0
27	1	0	0	0	0	1	0
49	1	1	0	0	0	0	0
29	1	0	0	1	0	0	0
30	1	1	0	0	0	0	0
31	1	0	0	0	0	0	1
32	1	1	0	0	0	0	0
43	1	0	0	0	0	0	1
58	1	0	0	0	0	1	0
44	1	1	0	0	0	0	0
24	1	1	0	0	0	0	0
28	1	1	0	0	0	0	0
33	1	1	0	0	0	0	0
34	1	1	0	0	0	0	0
26	1	1	0	0	0	0	0
62	2	0	0	0	0	2	0
35	2	0	0	0	0	2	0
36	2	0	0	1	0	0	1
38	2	0	0	2	0	0	0
39	2	0	0	0	0	2	0
40	2	0	0	1	0	1	0

User ID	(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(D) Viewed Evidence from Regular Page	(E) Viewed Evidence from Profile Page	(F) Used Highlight Tool	(G) Changed Profile attributes
42	2	0	0	1	0	0	1
51	2	0	0	1	0	1	0
54	2	0	0	2	0	0	0
37	2	1	0	1	0	0	0
41	2	0	0	0	0	0	2
45	3	0	0	2	0	1	0
92	3	1	0	1	0	1	0
47	3	0	0	2	0	1	0
52	3	1	0	0	0	0	2
55	3	0	0	0	0	1	2
57	3	0	1	2	0	0	0
70	3	2	0	0	1	0	0
61	3	0	0	3	0	0	0
46	3	2	0	0	0	1	0
48	3	0	0	0	0	1	2
50	3	2	0	0	0	1	0
53	3	0	1	1	0	1	0
59	3	2	0	0	0	1	0
60	3	2	0	1	0	0	0
63	4	0	0	0	0	3	1
84	4	0	0	2	0	2	0
65	4	0	0	0	0	0	4
66	4	0	0	0	0	0	4
69	4	0	0	1	0	1	2
56	4	1	0	1	0	0	2
72	4	0	0	0	0	0	4
73	4	1	0	3	0	0	0
68	4	1	0	1	0	2	0
71	4	2	0	1	0	1	0
64	4	0	0	2	0	0	2
85	5	1	1	2	0	0	1
75	5	1	0	1	0	0	3
76	5	0	0	5	0	0	0
82	5	0	0	1	0	0	4
77	5	1	0	0	0	0	4
78	5	0	0	0	0	5	0
74	5	2	0	0	0	3	0
93	6	1	0	1	0	0	4
88	6	0	0	1	0	3	2
81	6	1	0	2	0	1	2
79	6	1	1	2	0	0	2
86	6	3	0	0	2	1	0
80	6	3	0	0	0	0	3
83	7	1	1	2	0	0	3
87	7	1	1	1	0	4	0
97	8	3	0	2	1	0	2
90	8	2	0	4	0	1	1
100	9	1	0	0	0	1	7

User ID	(A) Total Scrutiny Actions	(B) Opened Profile from menu	I Opened Profile from evidence	(D) Viewed Evidence from Regular Page	(E) Viewed Evidence from Profile Page	(F) Used Highlight Tool	(G) Changed Profile attributes
96	9	2	0	0	2	1	4
91	9	5	0	0	0	4	0
94	9	1	1	5	0	0	2
99	10	3	0	0	0	0	7
89	11	4	0	0	0	0	7
98	12	0	2	2	0	2	6
101	12	2	2	2	4	2	0
95	13	0	0	0	2	0	11
105	14	3	2	2	3	0	4
102	16	5	0	3	0	1	7
103	17	2	3	7	0	5	0
104	22	3	0	4	2	8	5

8.4.4 Top Users of Scrutinisation Tools

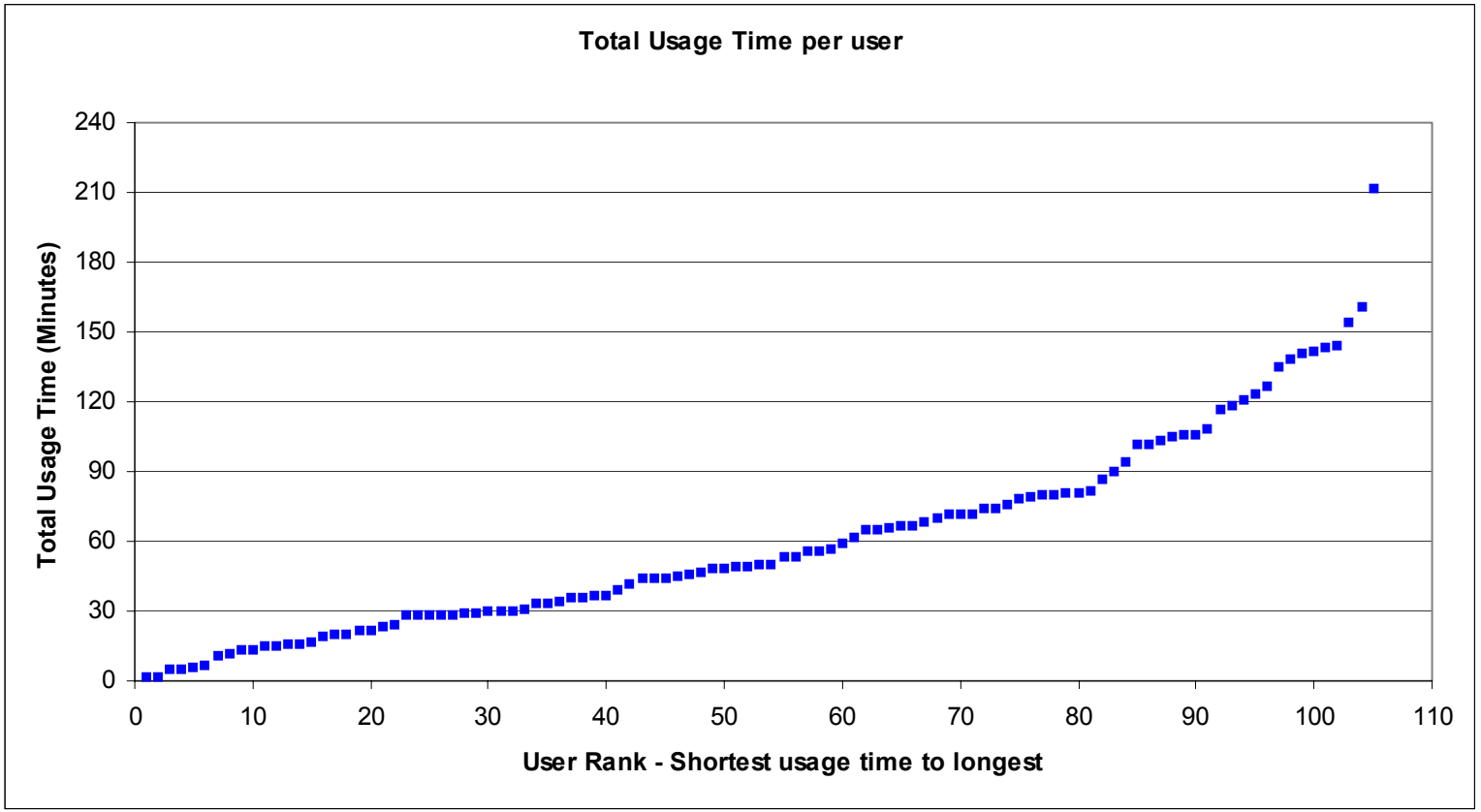
We studied the five users who made the most use of the scrutinisation tools. Their activity in SASY is summarised in Table 8.14. The accounts show these users seemed to get involved trying to understand and control the personalisation to better suit their needs.

Table 8.14 – Details of the scrutiny actions of the top 5 users.

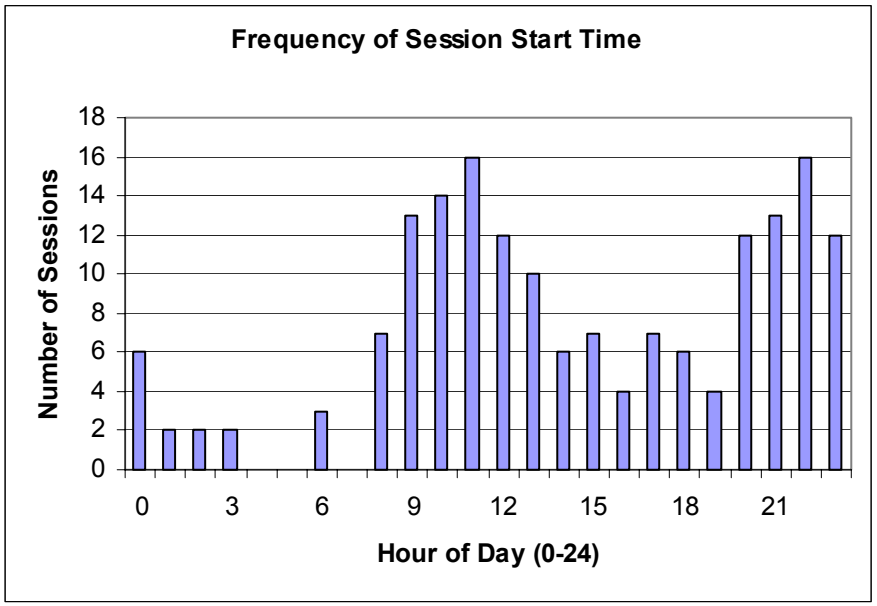
User ID	Total Scrutiny Actions	Pre-test Score	Post-test Score	Scrutiny Actions
95	13	4	9	• Before pre-test, navigated to profile directly and changed it to state the user had knowledge of UNIX and changed system preference to Highlight personalisation on each page by default. Soon after, used evidence tool to explore evidence about their knowledge and then changed profile back to original state. • Then completed pre-test and worked through the whole course with personalisation highlighted on each page by default. Completed post test.
105	14	9	10	• Completed pre-test but got one question wrong about chmod. Immediately, after pre-test, clicked for evidence about why SASY stated they knew nothing about the chmod command, then changed profile to state they had some knowledge about it. • Worked through course and completed post-test. Immediately after post-test, inspected coursemap, sought evidence for and changed their Goal attribute so that SASY would present a different exercise. • At the start of their second session, changed system preference to Highlight personalisation on each page by default but changed it back soon after. • At the end of their second session, sought evidence about all their knowledge attributes.
102	16	9	9	• Completed pre-test and worked through two modules.

User ID	Total Scrutiny Actions	Pre-test Score	Post-test Score	Scrutiny Actions
				Having completed the second module, used highlight tool on home page and attempted to access pre-test again. • Having completed post-test, accessed course map, sought evidence about knowledge and preferences and reset their quiz scores in the profile and started working through quiz questions again. • At the start of their second session, changed system preference to Highlight personalisation on each page by default and continued working through quiz questions. • Having finished the quizzes they inspected the course map and profile. Then changed their Goal attribute so that SASY would present a different exercise. Then inspected the profile.
103	17	8	10	• Completed pre-test, course material then post-test. Immediately after post test, explored coursemap, profile, highlight tool on home page to access post-test again. They either wanted to repeat the post-test or access their score to provide a print out as proof of completing their homework assignment. Then explored evidence about their knowledge of chmod. • Then sought evidence about their Goal attribute, the Assessed attribute and their file system knowledge.
104	22	10	9	• Completed pre-test, and Shell module. Then used highlight tool on home page. • Completed post test then sought evidence about why SASY thought they wanted to see advanced concepts. Inspected the profile and coursemap. • At the start of their second session, used the highlight tool on the home page to access the link to the File System content pages that had been removed by personalisation. They read through this material again. Then accessed the profile and sought evidence about their chmod knowledge. • Then changed their profile to remove jokes, hints, remove all quiz questions and change their Goal attribute. • At the end of the second session, used the highlight tool on the coursemap page.

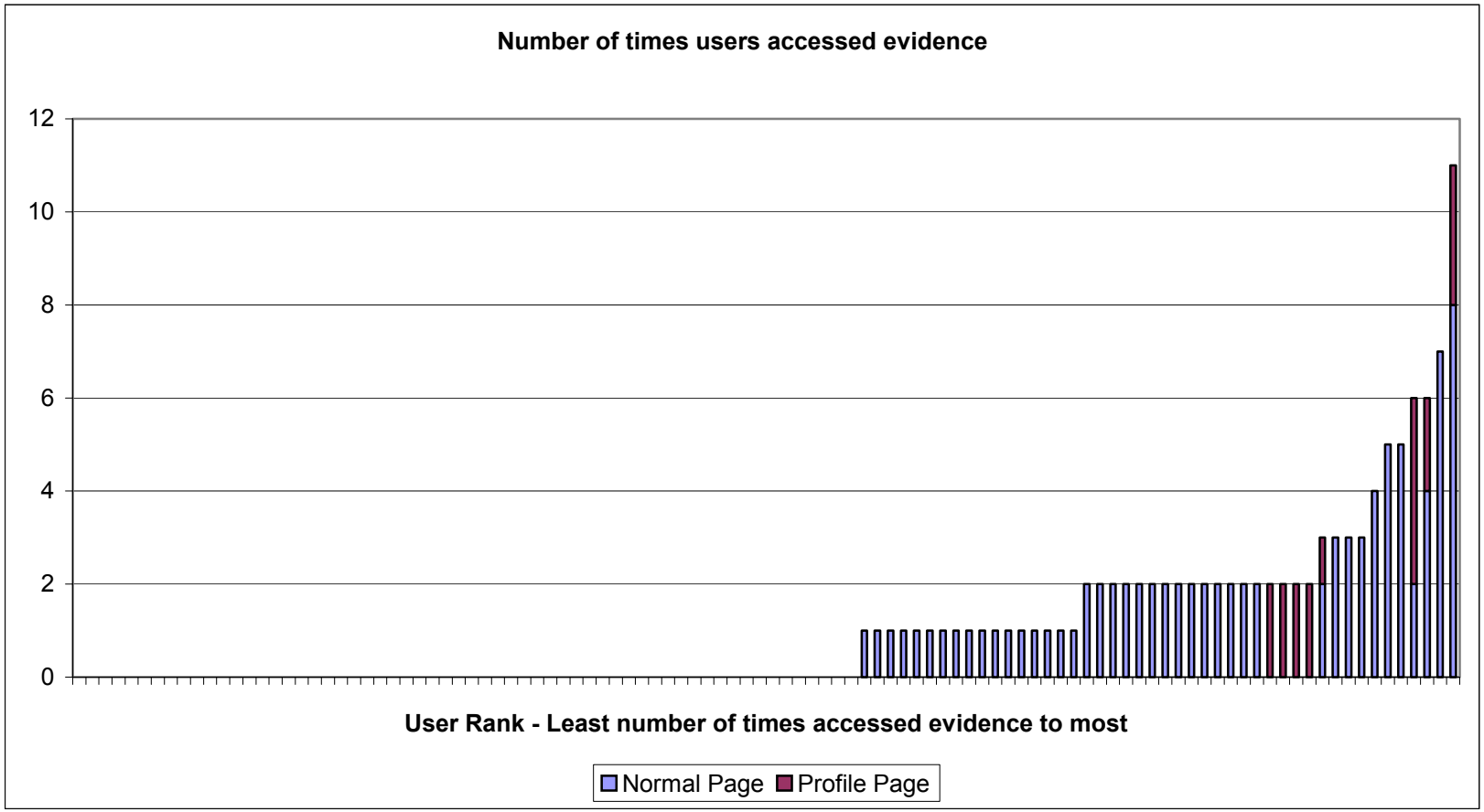
8.4.5 Graph of Total Usage Time per user



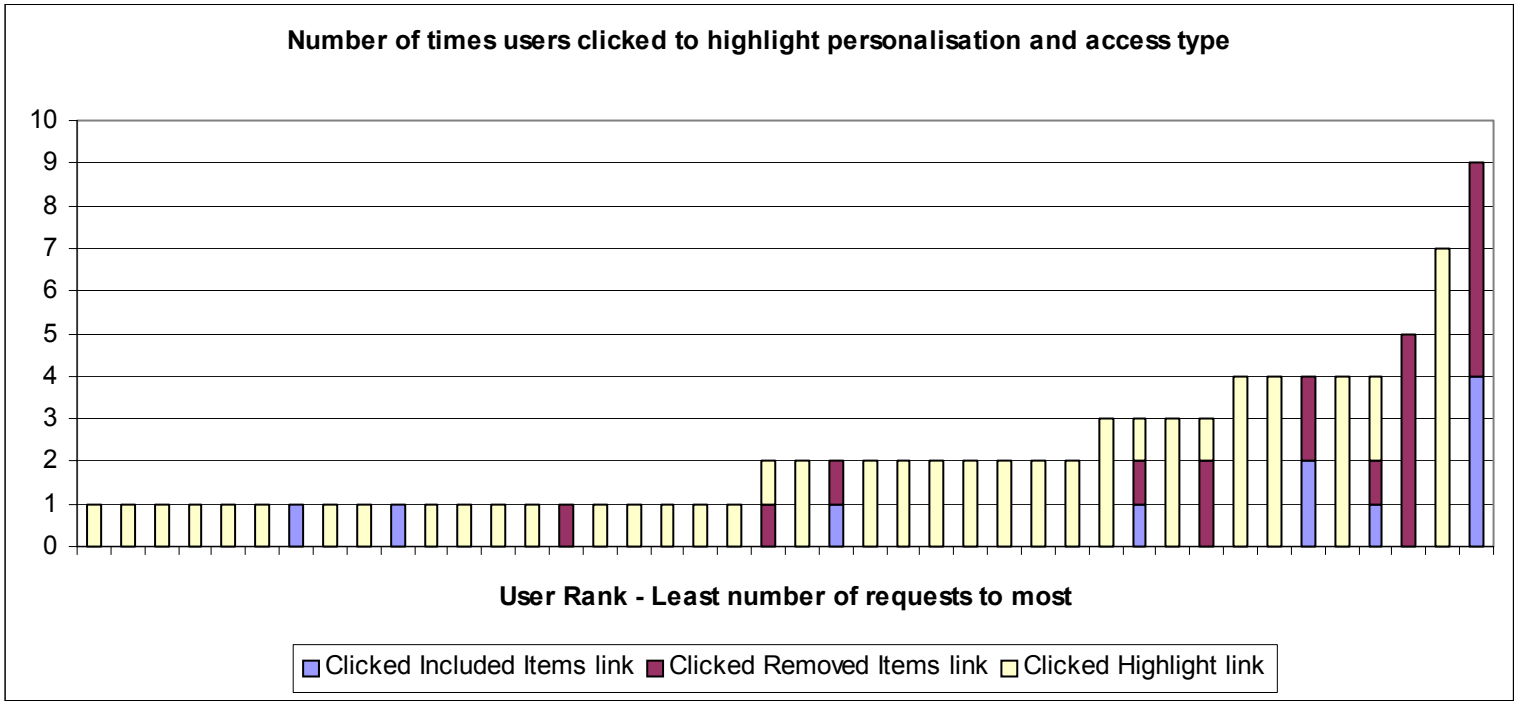
8.4.6 Frequency of Session Start Time



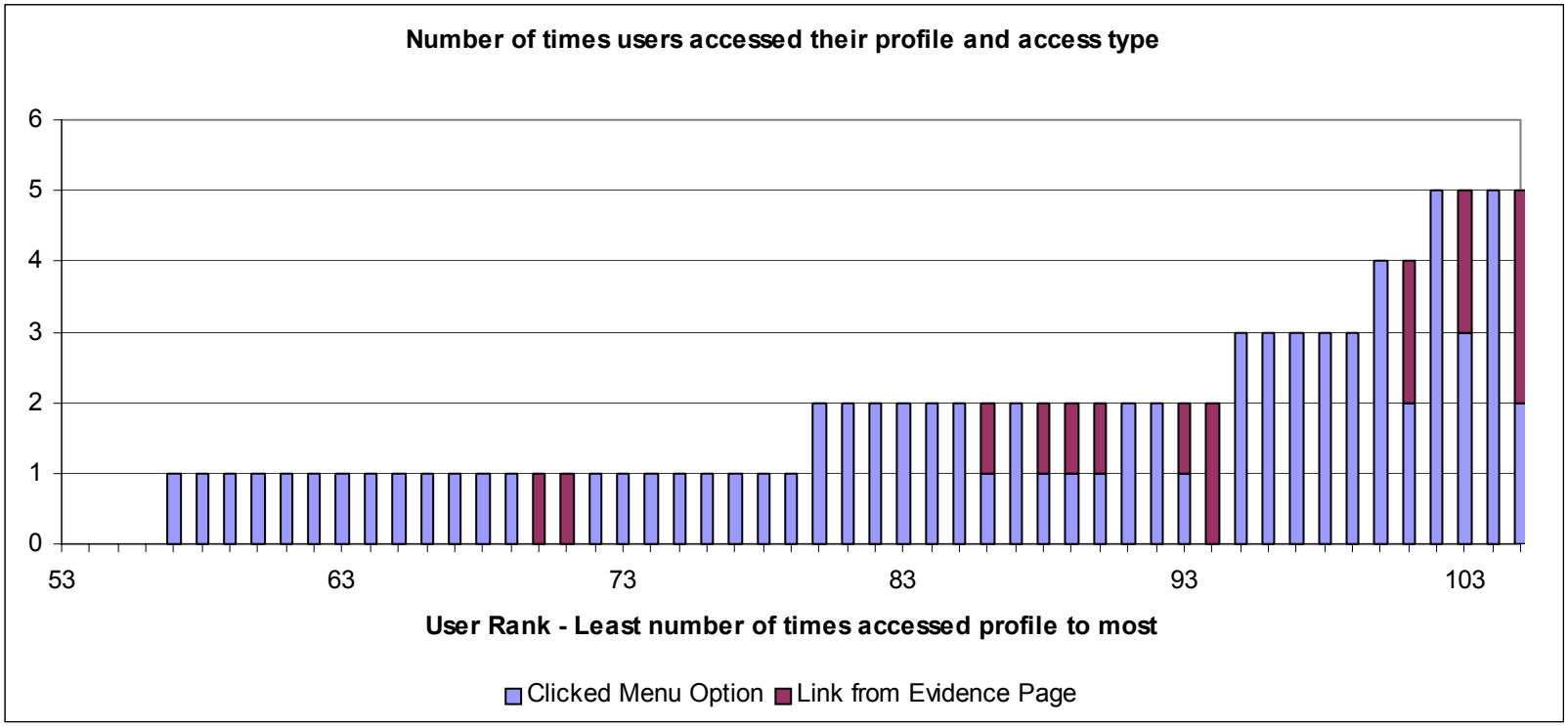
8.4.7 Number of times users used Evidence Tool and access type



8.4.8 Number of times users used Highlight Tool and access type



8.4.9 Number of times users used Profile Tool and access type



8.4.10 Email feedback from user 105 following their usage of SASY

Q: When you first logged in and completed the pre-test, you got an almost perfect score but SASY assumed you knew nothing about Chmod. You immediately clicked the link "You do not know about the chmod command" then accessed your profile and changed it so say "You know the chmod command but haven't passed the quiz". Would it be safe to assume you were surprised SASY said you did not know chmod and you wanted to find out why it said this and change the system's misconception about you?

User 105: "That is quite correct. I assumed that SASY would use the results of the pre-test to tailor the level of questions and hints, and I was surprised at the response regarding chmod."

Q: You also accessed the profile and the evidence about whether you knew Chmod, the File System and the Shell at the very end of your session. Was this to reflect on your learning, i.e. see what you had learnt, why the system now said you knew these topics?

User 105: "That is true, I did go back to see what I had learnt from SASY, and I was also curious about SASY and what it had determined about me. I also wanted to see if there was anything else I had missed. I chose to access this information at the end of the session largely because SASY gave feedback throughout the session as to what it knew about me, with the most useful information being which tests I had attempted and what I had yet to do. There was little need to see what the system had assumed about me until the end of the session, when curiosity took over."

8.5 Evaluation 2 – Qualitative Study

8.5.1 Results

Table 8.15 – Results of tasks for Evaluation 2.

Key:

Forms of scrutinisation:

A Understand what aspects of the hypertext were personalised. That is, what parts of the content were added to and/or removed from the user's personalised view of the page.

B Understand what aspects of the user model caused each personalisation. See the implications of changing their user model, in terms of how it effects personalisation. Be able to visualise what-if scenarios to understand the impact of changes to their profile without making any changes.

C View and understand evidence of how user model attributes were set, whether set directly by the user or inferred by the system.

D View or change aspects of their user model to control the personalisation.

Other labels used:

Used Hint = User clicked on Hint hyperlink to display the text "In SASY, look at the right hand side of the page in the personalisation column".

Int. = Evaluation observer intervention required to help user complete task.

? = User could not complete task

		Forms of Scrutinisation used by participants												
		Iteration 1			**2**		**3**				**4**	**5**		
Task	**Description**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
2	Examine personalised TV Guide page. Find out all the programs that play in this timeslot. *Expected scrutinisation: A*	D (Used Hint) A	A	D	A	A	C,D (Used Hint) A	A	A	A	A	A	A	A
3	Which programs are added for people who are interested in Current Affairs programs? *Expected scrutinisation: A +B or D*	D	D ?	D ?	A, B	A, B	C D	?	(Int.) A,B	D	(Int.) A,B	A, B	A, B	A, B
4	… Children's programs?	?	?	?	A, B	A, B	D A,B	?	A,B	?	A,B	A, B	A, B	A, B

		Forms of Scrutinisation used by participants												
		Iteration 1			2		3				4	5		
Task	**Description**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
5	… Lifestyle programs?	?	?	?	A, B	A, B	A,B	?	A,B	(Int.) A,B	A,B	A, B	A, B	A, B
6	Change the personalisation so that only Current Affairs programs are included in your 4:30 – 5:30 schedule. *Expected scrutinisation: D*	D	D	D	C, D	D	D	C,D	D	D	D	D	D	D
7	Change the personalisation so that the program Aerobics Oz Style is included in your schedule. *Expected scrutinisation: A +B + D or D*	D (Int.) A,B,D	D	D ?	D A,B,C,D	A,B,D	D A,B,D	A,B,C,D	A,B,D	A,B,D	A,B,D	D	D A, B	D (Int.) A,B,D
8	Change the personalisation so that the current affairs program that is included in your 6 – 7am schedule is the Today show. *Expected scrutinisation: A +B + D or D*	A,B,D	D	D	A,B, D	D	A,B,D	C ?	D	A,B ?	A,B,D	D	A,B,D	A,B,D
9	Access Program Index page and freely click on programs of interest. *(This task did not require the use of scrutinisation tools).*	-	-	-	-	-	-	-	-	-	-	-	-	
10	Return to TV Guide constructed earlier and explain: Why it has changed? Why programs have been added? Why system states you are a member of a special interest group interested in religious themes? *Expected scrutinisation: A +B + C*	A,B (int.) C	(int.) A,C	(int.) A,B C	A,B (int.) C	A,B,C	A,B,C	(int.) A,B C	A,B (int.) C	A,B,C	A,B,C	D (int.) A,B C	A,B (int.) C	A,B,C

		Forms of Scrutinisation used by participants												
		Iteration 1			2		3				4	5		
Task	**Description**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
11	Change the personalisation such that only Current Affairs programs are included in your 4:30am – 5:30am schedule, as before. *Expected scrutinisation: D*	D	D	D	D	D	D	D	D	D	D	D	D	D

Table 8.16 – Responses to Qualitative Survey for Evaluation 2.

		Participants												
		Iteration 1			2		3				4	5		
Question	**Description**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
1	I knew how to inspect the personalisation to see what items were included or removed from my personalised view.	Agree	Strongly Agree (once I knew how)	Strongly Agree (once I knew how)	Agree	Agree	Strongly Agree	Neutral (not obvious)	Agree	Agree	Strongly Agree	Agree	Agree	Agree
2	It was easy to inspect the personalisation to see what items were included or removed from my personalised view.	-	-	-	Agree	Agree	Strongly Agree	Neutral (not obvious)	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Agree	Agree

		Participants												
		Iteration 1			2		3				4	5		
Question	**Description**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**
3	I knew that by changing my answers to questions on the Profile page I could change what was included and removed by the personalisation.	Strongly Agree	Agree	Strongly Agree	Agree	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Agree	Agree
4	It was easy to change what was included and removed by the personalisation.	-	-	-	Strongly Agree	Agree	Agree	Neutral (confused by SIG attributes in profile)	Strongly Agree	Strongly Agree	Agree	Neutral (could not find Profile link in menu)	Agree	Agree
5	I feel it is useful to be able to inspect and control personalisation.	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Agree	Agree	Agree	Agree
6	Overall, I felt in control on the personalisation.	-	-	-	-	-	-	Agree	Strongly Agree	Agree	Agree	Strongly Agree	Agree	Strongly Agree
7	Overall, I understood how the personalisation worked.	-	-	-	-	-	-	Agree	Strongly Agree	Agree	Agree	Strongly Agree	Agree	Strongly Agree

8.5.2 Iteration 1 – User Observation Notes

User 1 – Observation Notes

Part 1: Guide to Personalisation in SASY
Found guide long and tedious

Part 2: TV Guide Topic

2) First accessed profile. Then clicked the hint hyperlink to display the text "In SASY, look at the right hand side of the page in the personalisation column". User then clicked the removed items link and answered the question correctly.

3-5) Was able to click items removed link and items added link to correctly indicate which programs were removed/added by personalisation. However, did not initially notice mouse over tool that explained why each item was added or removed. Changed profile to answer questions 3 rather than using highlight tool.

- Correctly changed profile easily.
- First tried to change profile by adding interest in sport but this did not help. Was stuck and needed help to continue. When asked by the instructor to examine interface carefully and read the instructions, user understood how to use the tool.

8) Correctly completed question easily.

10) Was able to use Highlight tool but could not find evidence tool without intervention

11) Correctly changed profile easily.

User clearly understood that profile stored information about them, that it drove the personalisation and that changing the profile could change personalisation.

Once user had some practice using the scrutinisation tools, they said it was easy to use them.

User commented the link "Click to highlight included/removed items on this page" in the personalisation panel was confusing. It was not clear what its purpose was and it was redundant because the included items and removed items links were sufficient to highlight the content included and removed b personalisation.

User 2 – Observation Notes

Part 1: Guide to Personalisation in SASY
User read introduction guide but did not seem to understand the information or how to use the tools. The user had difficulty completing the instructions that asked the user to practice using the tool.

Part 2: TV Guide Topic

2) Quickly found and accessed Highlight tool, but did not notice mouse over explanations.

3-5) Looked in profile, but could not complete because did not find mouse over explanations.

- Correctly changed profile easily.
- Looked in profile and guessed correctly to complete task.

8) Looked in profile and guessed correctly to complete task.

10) Used Highlight tool but still did not notice mouse over so did not know exactly why religious programs had been added. However, easily found evidence tool.

11) Correctly changed profile easily.

User clearly understood that profile stored information about them, that it drove the personalisation and that changing the profile could change personalisation.

However, user struggled to understand the Highlight tool because they did not notice the mouseovers text that explained by items were included or removed by personalisation.

To complete worksheet, user relied heavily on profile and guessing which profile attributes to change to change personalisation rather than using the Highlight and Evidence tools.

After the evaluation, the instructor explained how to use the tools at which point the user said it was easy to use the tools once you were shown how.

This user had a very low level of computer literacy and had difficulty switching between the worksheet and SASY browser windows.

User 3 – Observation Notes

Part 1: Guide to Personalisation in SASY
User read introduction guide but did not deeply understand the information or how to use the tools. The user expressed this was because the content was too "busy" and difficult to read and focus.

Part 2: TV Guide Topic

2) Looked in profile and did not notice Highlight tool so could not complete.

3-5) Looked in profile, but could not complete because did not find mouse over explanations.

- Correctly changed profile easily.
- Looked in profile and attempted to guess but could not complete task.

8) Looked in profile and guessed correctly to complete task.

10) Was stuck. When asked by the instructor to examine interface carefully and read the instructions, user understood how to use the tool. Also found evidence tool.

11) Correctly changed profile easily.

User clearly understood that profile stored information about them, that it drove the personalisation and that changing the profile could change personalisation.

This user firstly accessed the profile to change personalisation rather than using the Highlight or Evidence tools. User later said this was because the training material had discouraged them from using the tools because it made them sound so complicated.

When user was asked to explore interface, they eventually worked out how to use the Highlight and Evidence tools.

User suggested a number of changes to the working of the worksheet to make the questions clearer.

On discussion with the user following the experiment, they said the tools were actually very easy to use but the introduction over complicated their functionality, which had initially discouraged their use of the tools. User added "I didn't see the need for lengthy introductions" in regards to the online help material.

8.5.3 Iteration 2 – User Observation Notes

User 4 – Observation Notes

Part 1: Guide to Personalisation in SASY

On Page: 2.1 How the Personalisation Works:

- Moved mouse cursor over *items removed* link and said "This will tell me what was removed, then clicked it.
- Then moved mouse over highlighted items and read out the pop up information, i.e. "tofu was removed because You are not vegetarian".
- User also stated they understood the highlight tool.

On Page: 2.2 See what is personalised – Highlight Tool:

- Clicked *items included* link, page highlighted personalisation.
- Clicked *items removed* link, page returned to normal view.
- User was confused because Step 2 on the page referred to highlighted items on the page, but because the page was now in normal view, nothing was highlighted. User was being confused by the changing modality of the page.

On Page: 3. Control the personalisation

- User correctly guessed how changing the profile would change the personalisation, but did not use highlight tool to make prediction.

In all training pages, user answered Yes to the question "Do you feel confident you understand the information above?".

Part 2: TV Guide Topic

Q2) Clicked *items removed* link and correctly indicated which programs were removed by personalisation. Did not use any hints.

Q3) Clicked *items added* link and used mouse over tool to correctly indicate which programs were added for people interested in current affairs.

Q4) Page still had personalisation highlighted. User used mouse over tool to correctly indicate which programs were added for people interested in children's programs.

Q5) Page still had personalisation highlighted. User used mouse over tool to correctly indicate which programs were added for people interested in lifestyle programs.

Q6) At this point, user had forgotten how to access profile page. They did not see the link in the menu but remembered they could access it from the evidence pop-up. User accessed profile through pop up and removed interest from all categories other than current affairs.

Q7) Rather then using highlight tool, user opened profile and expressed interest in sport and lifestyle. When Aerobics did not appear in the TV Guide, user said "it wasn't either of those, so to find out why…" as they used Highlight tool and moved mouse over Aerobics to pop up the reason it was removed. User continued "so its because I'm not a member of SIG2. I don't know what SIG2 is. So I need to join SIG2, whatever that is.". User clicked on evidence they are non a member of SIG2, accessed and changed profile to correctly complete this task.

Q8) User was very confused about the wording of this worksheet question. User was confused by the mouse over explanation because they did not realise that there was nested adaptation. Admittedly, the pop up explanation did not explain that the programs were included because of an interest of current affairs. User spent some time to examine the differences between the pop up message for programs and eventually worked out the difference was the channel preference. User then access profile and specified their channel preference to correctly answer this question. User state "this is just a guess".

It seemed the user could not easily recognise which link on the page accessed the profile and which link accessed the personalisation highlight.

Q10) Was able to use highlight tool to determine which programs were added, but was not able to find out why SASY made them a member of SIG 2. That is, user could not recall that the evidence tool could help with this.

User was also getting mixed by between SIG 1 and SIG 2.

Q11) User access profile directly and removed interest from all categories except current affairs.

Part 3: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.
- Agreed

Q2) It was easy to inspect the personalisation to see what items were included or removed from my personalised view.
- Agreed

Q3) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.
- Agreed

Q4) It was easy to change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.
- Strongly Agreed

Q5) I feel it is useful to be able to inspect and control the personalisation
- Strongly Agreed

Additional Comments
Mentioned the Evidence Tool should have a heading in the training guide, because it is not obvious what the effect will be of clicking links on the personalisation column that list attributes from the profile. User suggested adding a label "Why?" for the evidence tool, similar to how www.amazon.com labels its link to seek an explanation of the recommended items.

User 5 – Observation Notes

Part 1: Guide to Personalisation in SASY

On Page: 2.1 How the Personalisation Works:
- Clicked removed items link, and used mouse over tool carefully. Then clicked added items link. Suggested "I expected the two links would show a different page".

On Page: 3. Control the personalisation
- User correctly guessed how changing the profile would change the personalisation, using the highlight tool to make the prediction.
- User confidently changed the profile and stated this produced expected results.

In all training pages, user answered Yes to the question "Do you feel confident you understand the information above?".

Part 2: TV Guide Topic

Q2) User almost immediately clicked *items removed* link and *items added* links and correctly indicated which programs were removed by personalisation. Did not use any hints. User copy/pasted the SASY page to the worksheet – indicating a high level of computer literacy.

Q3) Clicked *items added* link and used mouse over tool to correctly indicate which programs were added for people interested in current affairs. However, required about 30 seconds to complete this task and stated "I'd forgotten about the pop up. It's something you just have to learn about or get reminded." User eventually read instructions on the page and found out about the mouse over function and completed the task.

Q4) User used mouse over tool to correctly indicate which programs were added for people interested in children's programs and also viewed evidence about this attribute.

Q5) Page still had personalisation highlighted. User used mouse over tool to confidently to indicate which programs were added for people interested in lifestyle programs. However, user spent some time wondering what SIG1 meant and said "assume the SIG1 programs are not life style programs but I am not sure".
User did not think to click on evidence link at this point.

Q6) User accessed profile and removed interest from all categories other than current affairs. User also noticed SIG1 and read this, stating they finally knew what SIG1 meant.

Q7) User immediately access removed items link and used mouse over tool to see that the Aerobics program was not shown and said it was "because I'm not a member of SIG2". User then correctly updated profile to make them selves a member of SIG2.

Q8) User quite easily answered this question. They clicked the included items link and examined the included programs and compared the different explanations for each to determine that the channel preference was different for each. User then accessed profile and correctly updated channel preference attribute..

Q10) User examined several programs and returned to their TV Guide to find religious programs had been added. User stated "Oh my God! I am a member of SIG1. I thought I became a member of SIG2?" User used evidence tool for SIG1 to see why they had been added to SIG 1, read the reason and laughed. User clearly understood how to use the highlight and evidence tools at this point.

Q11) User access profile directly and carefully removed interest from all categories except current affairs.

Part 3: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Agreed

Q2) It was easy to inspect the personalisation to see what items were included or removed from my personalised view.

- Agreed

Q3) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Agreed

Q4) It was easy to change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Agreed

Q5) I feel it is useful to be able to inspect and control the personalisation

- Agreed

Additional Comments
User said they were surprised and upset the system included religious programs without first asking the user if this was ok.

User's self-description of their computer literacy:
"Very reasonable, spent most of my career in computing, usage mostly at the GUI level. But if something breaks I'm not confident I can fix it."

User performed well in training, but initially forgot about evidence and mouse over tool. After some practice the user became very proficient in using all tools.

8.5.4 Iteration 3 – User Observation Notes

User 6 – Observation Notes

Part 1: TV Guide Topic

2) After about 40 seconds of inspecting the SASY page, User retuned to the worksheet and clicked the hint hyperlink to display the text "In SASY, look at the right hand side of the page in the personalisation column". User then clicked the removed items link and answered the question correctly. User copy/pasted the programs in to the worksheet.

User did not notice mouse over tool.

3) User first clicked evidence link "You are not Interested in Current Affairs". Read evidence pop-up and clicked through to profile. User read through profile page and changed their profile to include news items. User then examined the TV Guide page and noted the programs that had been included. The user thus answered the question correctly by using the profile instead of the highlight tool.

4) User firstly used the highlight tool, but did not notice mouse over so could not answer question. User then navigated to profile through the evidence pop up for "You are Interested in Current Affairs" and changed their profile back to specify no interest in current affairs. 10 seconds after returning to the TV Guide the user again clicked the removed items link. About 20 seconds later the user noticed the mouse over tool and was able to identify the children's program using the mouse over tool.

- Page still had personalisation highlighted. User immediately used mouse over tool to confidently to indicate which programs were added for people interested in lifestyle programs. User answered the question in 12 seconds.

- After 20 seconds the user navigated to profile through the evidence pop up for "You are not Interested in Current Affairs". User confidently changed profile at add interest in Current Affairs and remove interest in Lifestyle programs.

- User first accessed profile page through the menu, but could not determine which profile attribute to change.

User returned to TV Guide and clicked removed items link and used mouse over tool to determine that Aerobics was not shown because the user is "Not a member of SIG 2". It took the user 10 seconds to use the mouse-over tool. User then accessed profile page and immediately changed the SIG 2 profile attribute to correctly complete the task.

8) User quite easily answered this question. User immediately clicked the included items link, examined the included programs and compared the different mouse over explanations (for about 20 seconds). User correctly determined that channel preference had to be specified as channel 9 to complete the task, then accessed profile through menu and correctly updated channel preference attribute.

9) User examined three programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User clicked the include items link to correctly indicate the programs that had been added. User said they were "Added because I'm now in SIG1".

User then clicked the evidence link to see why they had been made a member of SIG 1.

11) User access profile through menu and removed interest from all categories except current affairs to complete task in 10 seconds.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Strongly Agreed
- Comments: It seems that whether you click "items removed" or "items added" you'll get the same page. Being able to hover over an item and see what "rules" were applied was useful.

Q2) It was easy to inspect the personalisation to see what items were included or removed from my personalised view.

- Strongly Agreed
- Comments: Being able to hover over an item and see what "rules" were applied was useful.

Q3) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Agreed
- Comments: At first I was getting to the personalisation page by clicking "Why?" and then using the link on the pop-up to get there. After a few times I realised there was a "Your Profile" link at the top that took me to the same page – I probably would have used this sooner if I wasn't following this evaluation guide.

Q4) It was easy to change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Agreed

Q5) I feel it is useful to be able to inspect and control the personalisation

- Strongly Agreed
- Comments: Especially true for being able to change the automatic special interest group allocation.

Additional Comments
If this were a real TV Guide personalisation system I'd like to be able to give a grade of preference (1-5,etc) rather than interested/not-interested.
I guess this would complicate the explanation part of the system though ...

User 7 – Observation Notes

Part 1: TV Guide Topic

2) User quickly scanned the page and was observed to quickly move their gaze to the Personalisation panel on the page and examine it for about 15 seconds. User clicked removed items link and correctly determined which programs had been removed by personalisation. User did not notice mouse over tool. In fact, when the highlight mode was invoked the user quickly scrolled the page down, effectively hiding the instructions. Although the user moved the mouse over the coloured sections and the pop-up appeared, it was quickly closed when the user moved the cursor away (analogous to a tool tip).

3) The user correctly guessed the answer to the question, judging by the program names.

4) The user correctly guessed the answer to the question, judging by the program names.

- User clicked the evidence link for "You are interested in Lifestyle programs" but this did not help answer the question.

The user correctly guessed the answer to the question judging by the program names, but could not answer confidently.

- User clicked the evidence link for "You are interested in Lifestyle programs" because they had noticed (in the previous task) the pop-up had a link to the profile page. User accessed profile and confidently changed profile to reflect an interest in only Current Affairs programs.

- User removed items link confidently. However, as user not noticed mouse over tool was stuck at this point and said "I don't know".

The instructor asked the user to read the Instructions at the top of the page. User read the instructions and immediately said "Ah Ha" as if to indicate understanding. User then used mouse over tool over the Aerobics programs and said "What is SIG2?". User then accessed profile through the evidence pop up but by mistake changed SIG 1 attribute rather than SIG 2. On return to the page user realised they had changed the wrong attribute so they returned to the profile again and changed the SIG 2 attribute, correctly completing the task.

8) User did not think to use highlight or mouse over tool. Instead user clicked evidence link but this did not help complete the task. Eventually user correctly guessed the answer judging by the program names.

9) User examined three programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User examined page for and worksheet question for about 2.5 minutes and correctly noticed (by observation) which programs had been added to the TV Guide.

However, the user did not know why the programs were added, stating "I have no bloody idea. Because of membership in a SIG group? Because I set channel preference?"

Instructor: "How would you find out why the additional programs were added?"
User: "Profile?"
Instructor: "Is there anything else you can use?"

Then user clicked included items link and used mouse over for about 1 minute.
User then user the mouse over tool and realised that their membership in SIG 1 caused the program to be added when user manually selected membership in SIG 1 (in Q7). User commented "Oh Ok, that's what I thought, it's because of SIG 1. It should have already told me without making me search for it."

11) User access profile through evidence page and were recalling what they had set originally. User was not sure about the SIG 1 attributes and said "I think just SIG 1 is all I need to change, but we will check in a minute". This statement implied the user could confidently check the results of having changed their profile.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Neutral

Comments: *I did not know every time. It was not obvious [what was added/removed]. I needed to fiddle around to find the answers.*

Q2) It was easy to inspect the personalisation to see what items were included or removed from my personalised view.

- Neutral

Q3) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Agreed

Comments: *Not always because I did not know what SIG 1 and SIG 2 meant.*

Q4) It was easy to change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Neutral

Q5) I feel it is useful to be able to inspect and control the personalisation

- Agreed

Q6) Overall, I felt I could control the personalisation.

- Agreed

Q7) Overall, I understood how the personalisation worked.

- Agreed

Additional Comments
User commented about how the profile was initially constructed:
"I shouldn't have to pick preferences. It should learn about what programs I like to watch as I use the system. I shouldn't have to tell it."

User 8 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2) User looked at the SASY page and quickly found and clicked the removed link and correctly determined which programs had been removed by personalisation. User typed answers in to worksheet.

User did not notice the mouse over tool and did not read the instructions on the page as to how to use it.

3) Since user did not notice the mouseover tool, they initially did not know how to answer the question. They clicked the evidence for "You are not interested in Current Affairs programs" but this did not help complete the task. After a couple of minutes the user said "It doesn't say".

The instructor prompted the user to read the instructions on the SASY page. The user read the instructions that described the mouseover tool and used the mouse over tool to correctly indicate programs removed because of no interest in Current Affairs.

4) User correctly answered the question using the mouse over tool.

- User correctly answered the question using the mouse over tool.

- User was not sure whether to click "Your Profile" or click the evidence link. After some hesitation, user clicked the evidence link and from there accessed the profile. User correctly indicated interest in Current Affairs to complete the task.

- Confidently, user removed items link and found the Aerobics program (which had been removed) and used the mouse over tool, stating "because you are not a member of SIG

2". User then confidently changed profile to indicate an interest in SIG 2 and commented "ok lets change it…ok I am now a member of SIG 2".

The user had noticed and read the explanation of SIG 2 on the profile page and so was probably comfortable with the use of the code name "SIG 2".

8) User did not how to complete this task initially. First the user accessed the profile and read each attribute carefully. Here the user noticed the channel preference attribute and changed it. Although this correctly answered the question, it was a guess. The user did not think to use the highlight and mouse over tools here.

9) User examined three programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User clicked added items link and correctly indicated which programs had been added to the TV Guide. User stated they suspected that "by looking at programs in the TV Index the system might change their profile to include additional interests." The user was not surprised when new programs were added to their TV Guide.

The instructor asked, "what would you do to find out why these programs were added?"

The user first opened the profile, but did not realise they could access the evidence information from the profile. The user then returned to the page and clicked the evidence link. The user commented: "Because it has recognised that I have chosen a program that has religious and eligious themes to it and has included other programs of the same nature. By going to the TV guide and looking at programs I was interested in it recognised the program had a religious aspect and include other shows of that nature. It updated my profile."

11) User accessed profile from menu and correctly reset profile back to original state.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Agree

comments: Initially I wasn't totally confident on what i was doing, but after a couple of attempts I could navigate around the system easily

Q2) It was easy to inspect the personalisation to see what items were included or removed from my personalised view.

- Strongly Agreed

Q3) I knew that I could change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Strongly Agreed

Q4) It was easy to change the beliefs on the "Your Profile" page to change what was included and removed by the personalisation.

- Strongly Agreed

Q5) I feel it is useful to be able to inspect and control the personalisation

- Strongly Agreed

Q6) Overall, I felt I could control the personalisation.

- Strongly Agreed

Q7) Overall, I understood how the personalisation worked.

- Strongly Agreed

Additional Comments
User commented it would be nice if the user could somehow click and drag programs to a trash can icon to remove programs they were not interested in, rather than having to go to the profile which required a lot more effort.

The instructions are not obvious and you don't notice them easily. Perhaps you need to make it stand out more.

Computer literacy;
Had played with programming at High School but since then use computers mostly for data entry.

User 9 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2) First user clicked profile page and read it carefully, noticing the SIG 1 and 2 groups and stated he understood he was not a member of these "clusters" but could become one.

User then clicked removed items link but did not see instructions about the mouse over tool. User correctly determined all the programs available on this page.

3) To determine programs added for people interested in Current Affairs, user accessed profile and specified this interest rather than using highlight tool. Note that at this point user did not know the highlight tool could be used for this purpose.

4) Guessed answer.

- Could not answer question without taking a guess. Instructor asked participant to read the instructions about the mouse over tool. The user read and seemed to understand the mouse over explanations, then completed the task correctly.

- Confidently changed profile to correctly answer question.

- User clicked removed items link. After about a minute of inspecting the page and the mouse over tool, user was able to complete the task commenting "so I have to change it [the profile] so that I am a member of SIG2".

8) User immediately clicked removed items link and proceeded to use mouse over tool. However, because the user had no interest in Current Affairs, the mouse over was not sufficient to determine what to do to make only the Today show included.
[This is a limitation of the mouse over tool.]

User could not answer question, however, his process in attempting to answer it was what we would have hoped for, i.e. using the mouse over tool.

9) User examined three programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User asked "do I have to research to answer or can I guess". Instructor asked participant to research the answer. User clicked items added link and used mouse over tool to determine additional programs were added because the user had an interest in SIG 1. User the confidently used the evidence tool and read the explanation as stated by the system "You have been included in special interest group 1 (SIG1) which covers religious themes because you are interested in The Movie Show, this week features the movie Life and Times of Jesus Christ."

11) User access profile through menu and removed interest from all categories except current affairs to complete task in 10 seconds.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Strongly Agree

comments: Pop up boxes assisted in modification of personalisation preferences.

Q2) It was **easy** to inspect the personalisation to see what items were included or removed from my personalised view

- Strongly Agreed

Q3) I knew that by changing my answers to questions on the "Profile" page I could change what was included and removed by the personalisation.

- Strongly Agreed

Q4) It was **easy** to change what was included and removed by the personalisation.

- Strongly Agreed

Q5) I feel it is useful to be able to inspect and control personalisation.

- Strongly Agreed

comments: By having a strong degree of control over personalisation it assisted in getting the programs I wanted on TV with the minimum of fuss. I also did not want some of the programs recommended and it was also good to alter the personalisation to reflect this.

Q6) Overall, I felt I could control the personalisation.

- Agreed

comments: I think that I knew more about the SIG1/2 setup this would have assisted in getting more control over the programs I wanted to watch.

Q7) Overall, I understood how the personalisation worked.

- Agreed

comments: Once I understood how the system could be manipulated I was able to gain more control over the programs I wanted to watch.

Additional Comments

This was an interesting experiment. The answers were not apparent but relied partially on knowing on the popup boxes and partially on how to manipulate the system.

8.5.5 Iteration 4 – User Observation Notes

User 10 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2) User quickly found and clicked removed items link. The new window (added for SASY 3) popped-up, showing the instructions for the mouse over tool. User first looked at the Key, read it out and said "ok, that makes sense". User ignored the instructions and un-checked the checkbox "Display this Tip next time".

On closing the new pop-up Instructions window, user had not read the instructions and did not know about the mouse over tool.

3) User could not complete task 3 until the instructor prompted them to read the instructions. Having read the instructions, user demonstrated an understanding of the Highlight and mouse-over tool by using it to complete the task. User said "I saw this [mouse over popup] earlier but did not pay attention to it. I don't read web pages in detail so it was not obvious that this feature is available in the system. I would prefer to see all the explanations in one go, rather than after a mouse over. It is hard to remember what the explanation said after it disappears."

4) User easily used the mouse over tool now that they had learnt how to use it. User commented "I suppose I will use the trick I learnt", meaning they would use the mouse-over tool to get the answer.

- User easily used the mouse over tool now that they had learnt how to use it.
- Quickly found and clicked link to profile in the menu. Confidently changed profile, removing interest from all genres except Current Affairs, to correctly answer question.
- User verbalised their thought process as they completed the task (but was not asked to do so). "So I have to look at what's removed", user clicked removed items link and used mouse over tool to see why Aerobics had been removed. "Oh, I'm have to be a member of SIG1! So I go to my profile and make myself a member of SIG1, but I am not a health conscious person."

User easily completed this task.

8) User immediately clicked removed items link and proceeded to use mouse over tool. User noticed that the channel preference attribute was the factor differentiating the current affairs programs and said "Ah ha. I know what to do. I have to choose channel 9 and current affairs."

User easily completed task.

9) User examined four programs and returned to their TV Guide, where religious programs had been added by SASY in the background. As the user did this they commented "System gonna try to learn about me, but it may not have the full info it wants, because I might be interested in programs I didn't click on."

10) User clicked items added link and used mouse over tool to determine additional programs were added because the user had an interest in SIG 1.
To find out why they had been made a member of SIG1, user first looked in profile but did not notice the Evidence link there.
User returned to TV schedule page and eventually found evidence tool, which they understood.

11) User access profile through menu and removed interest from all categories except current affairs to complete task Easily.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Strongly Agree

comments: always on the right side of screen. But without the mouse over instructions, which were initially hard to find, the highlighting is useless because it does not tell you why."

Q2) It was **easy** to inspect the personalisation to see what items were included or removed from my personalised view

- Strongly Agreed

Comments: "even had colours to distinguish"

Q3) I knew that by changing my answers to questions on the "Profile" page I could change what was included and removed by the personalisation.

- Strongly Agreed

Q4) It was **easy** to change what was included and removed by the personalisation.

- Agreed

Q5) I feel it is useful to be able to inspect and control personalisation.

- Agreed

comments: "I feel that it may make very strong assumptions especially when I am interested in Gateaway in general, it made an assumption that I was interested in all programs that had eligious content. I may just happened to be interested in what Gateaway was offering at that point in time like cathedrals but I may not be interested in Jesus in the movie show."

Q6) Overall, I felt I could control the personalisation.

- Agreed

comments: to a certain extent, for the same reason as in Q5.

Q7) Overall, I understood how the personalisation worked.

- Agreed

Computer literacy:
I don't think the experiment needed a lot of technical competency.

8.5.6 Iteration 5 – User Observation Notes

User 11 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2) User quickly found and clicked removed items link. User easily found the explanations of why items were added and removed on the page.

3) User easily completed task by reading explanations in highlighted sections on the page.

4, 5) Same as for Q3.

- User knew they had to access profile but did not know how to get to the profile. First user tried clicking back button but when this failed they searched the interface carefully and

found the link in the menu at the top of the page. User easily changed profile to complete task.

- User first accessed profile. But on profile page, did not know whether the Aerobics program would be classified as lifestyle or sport. User first tried lifestyle but this was not correct. User returned to profile and noticed the special interest group section and made themself a member of SIG2. Effectively, user corrected guessed the task.

8) User immediately accessed profile and guessed the answer.

9) User examined three programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User first tried to access profile, but realised this would not help answer the question. Instructor asked "Is there anything on the TV Guide page that could tell you what programs were added and removed?"

The user then found and clicked the "Added items link", read the explanation to realise they had been made a member of SIG1. It seems the user had forgotten about the Highlight tool altogether and tried to access the profile first in every task.

In contrast, User quickly found the evidence link and clicked to learn why SASY had made the user a member of SIG1. The user noticed the evidence pop up had a hyperlink to the profile and said "that's good!".

11) User access profile through menu and removed interest from all categories except current affairs to complete task. User did this easily.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.

- Agree

comments: "Took some time to work it out – but not too long" This comment is interesting because the user found it quickly at the beginning of the experiment forgot about it later on, and then struggled to find it.

Q2) It was **easy** to inspect the personalisation to see what items were included or removed from my personalised view

- Agreed

Q3) I knew that by changing my answers to questions on the "Profile" page I could change what was included and removed by the personalisation.

- Agreed

Q4) It was **easy** to change what was included and removed by the personalisation.

- Neutral.

Comments: "The navigation was not so obvious at first, but then I found the menu."

Q5) I feel it is useful to be able to inspect and control personalisation.

- Agreed

Q6) Overall, I felt I could control the personalisation.

- Strongly Agreed

Q7) Overall, I understood how the personalisation worked.

- Strongly Agreed

Additional Comments:

"Challenging and interesting"
Suggested reversing the highlight colour coding as the yellow used to highlight removed items "was too bright".

User 12 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2-5) Same as for user 11.

- User knew they had to access profile. User examined TV Guide page and scrolled up to find link to profile in menu.

User accessed and confidently changed profile to complete task.

- User first accessed profile. But on profile page, did not know which profile element to change. Instructor asked participant if there was a way to tell why the program had not been added to the TV Guide schedule on the TV Guide page, rather than the profile. User said "oh ok, I suppose I can click the removed link". User clicked the removed items link and used explanation displayed by highlight tool to complete task.

8) User immediately clicked added items link and used mouse over explanations to correctly complete task.

9) User examined programs and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User quickly clicked added items link to determine what programs had been added and read the explanation that they were now a member of SIG 1.

Initially user did not see the evidence link, but after some exploration of the interface user was able to answer question.

11) User access profile through menu and removed interest from all categories except current affairs to complete task. User did this easily.

Part 3: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.
- Agreed

Q2) It was **easy** to inspect the personalisation to see what items were included or removed from my personalised view
- Agreed

Q3) I knew that by changing my answers to questions on the "Profile" page I could change what was included and removed by the personalisation.
- Agreed

Q4) It was **easy** to change what was included and removed by the personalisation.
- Agreed

Q5) I feel it is useful to be able to inspect and control personalisation.
- Agreed

Q6) Overall, I felt I could control the personalisation.
- Agreed

Q7) Overall, I understood how the personalisation worked.
- Agreed

Additional Comments:
User commented that the Added items link should be on top in the right hand side of the page, not the Removed items link.

User 13 – Observation Notes

User did not do training.

Part 1: TV Guide Topic

2-5) Same as for user 11.

- User quickly found link to profile in menu. User accessed and confidently changed profile to complete task.

- User first accessed profile and said "Is it lifestyle?". On profile page, did not know which profile element to change.

Instructor asked participant if there was a way to tell why the program had not been added to the TV Guide schedule on the TV Guide page, rather than the profile. User clicked the removed items link and used explanation displayed by highlight tool to complete task.

8) User immediately clicked added items link and used mouse over explanations to correctly complete task.

9) User examined one program and returned to their TV Guide, where religious programs had been added by SASY in the background.

10) User quickly clicked added items link to determine what programs had been added and read the explanation that they were now a member of SIG 1.

Initially user did not see the evidence link, but after some exploration of the interface user was able to answer question.

11) User access profile through menu and removed interest from all categories except current affairs to complete task. User did this easily. User laughed when they had read the description of why SASY had made them a member of the religious special interest group.

Part 2: Qualitative Feedback

Q1) I knew how to inspect the personalisation to see what items were included or removed from my personalised view.
- Agreed

Q2) It was **easy** to inspect the personalisation to see what items were included or removed from my personalised view
- Agreed

Q3) I knew that by changing my answers to questions on the "Profile" page I could change what was included and removed by the personalisation.
- Strongly Agreed

Q4) It was **easy** to change what was included and removed by the personalisation.
- Strongly Agreed

Q5) I feel it is useful to be able to inspect and control personalisation.
- Agreed

Q6) Overall, I felt I could control the personalisation.
- Strongly Agreed

Q7) Overall, I understood how the personalisation worked.
- Strongly Agreed

Additional Comments:
User suggested making interests in the profile checkboxes rather than radio buttons as this might be easier to read.

Chapter 9 Bibliography

Ackerman, M.S.; Cranor, L.F. and Reagle, J. (1999): Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences. In Proceedings, ACM Conference on Electronic Commerce, Nov. 1999, pp. 1-8.

Alpert, S.R., Karat, J., Karat, C-M., Brodie, C., & Vergo, J.G. (2003). User Attitudes Regarding a User-Adaptive eCommerce Web Site. User Modeling and User-Adapted Interaction, 13(4), pp. 373-396.

Arteaga, C., Fabregat, R. (2004) Adaptive Support for Collaborative and Individual Learning (ASCIL): Integrating AHA! and CLAROLINE. In proceedings: De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer, pp. 279 – 282.

Baudisch, P., Bruecker, L. (2002) TV Scout: Lowering the Entry Barrier to Personalized TV Program Recommendation. In Proceedings: De Bra, P., Brusilovsky, P., Conejo, R. (Eds.) (2002) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Second International Conference, AH2002, Malaga, Spain, May 29-31, 2002. Lecture Notes In Computer Science, 2347, Springer.

Bicking, I. (2003) The Web Framework Shootout, http://www.python.org/pycon/papers/framework/web.html, accessed August 2003.

Bohnenberger, T., Jacobs, O., Jameson, A., & Aslan, I. (2005) Decision-theoretic planning meets user requirements: Enhancements and studies of an intelligent shopping guide. In H. Gellersen, R. Want, & A. Schmidt (Eds.), Pervasive computing: Third international conference, Springer, 2005, pp. 279-296.

Bontcheva, K., Wilks, Y. (2005) Tailoring Automatically Generated Hypertext. User Modeling and User-Adapted Interaction, Springer, 15, pp. 135 –168.

Brna, P., Self, J., Bull, S. & Pain, H. (1999). Negotiated Collaborative Assessment through Collaborative Student Modelling, *Proceedings of Workshop on Open, Interactive and Other Overt Approaches to Learner Modelling*, International Conference on Artificial Intelligence in Education 1999, pp. 35-42.

Brusilovsky, P. (1996) Methods and techniques of adaptive hypermedia.

User Modeling and User Adapted Interaction, Special issue on adaptive hypertext and Hypermedia, v 6, n 2-3, pp 87-129.

Brusilovsky, P. (2000) Adaptive Hypermedia: From Intelligent Tutoring Systems to Web-Based Education. In: G. Gauthier, C. Frasson and K. VanLehn (eds.) Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol. 1839, (Proceedings of 5th International Conference on Intelligent Tutoring Systems (ITS'2000), Montreal, Canada, June 21-23, 2000) Berlin: Springer-Verlag, pp. 1-7.

Brusilovsky, P. (2001) Adaptive hypermedia. User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), Kluwer Academic Publishers, pp. 87-110.

Brusilovsky, P., Maybury, M. (2002) From adaptive hypermedia to the adaptive Web, Communications of the ACM, Volume 45, Number 5 (2002), pp 30-33.

Brusilovsky, P., Sosnovsky, S. & Scherbinina, O. (2004). QuizGuide: Increasing the Educational Value of Individualized Self-Assessment Quizzes with Adaptive Navigation Support, in J. Nall & R. Robson (eds), *Proceedings of World Conference on E-Learning*, Association for the Advancement of Computing in Education, pp. 1806-1813.

Bull, S. and Pain, H. (1995). Did I say what I think I said, and do you agree with me?: inspecting and questioning the Student Model, Proceedings of World Conference on Artificial Intelligence in Education, AACE, Washington DC, USA, pp. 501-508.

Bull, S., Brna, P. and Pain, H. (1995b). Extending the scope of the student model. User Modeling and User-Adapted Interaction, 5(1), pp. 45-65.

Bull, S. & McEvoy, A. (2003). An Intelligent Learning Environment with an Open Learner Model for the Desktop PC and Pocket PC, in U. Hoppe, F. Verdejo & J. Kay (eds), *Artificial Intelligence in Education*, IOS Press, Amsterdam, pp. 389-391.

Bull, S., Brna, P. and Dimitrova, V. (Eds.) (2003). Proceedings of Learner Modelling for Reflection Workshop, Volume V of the AIED2003 Supplementary Proceedings, pp. 209-218.

Bull, S. & McKay, M. (2004). An Open Learner Model for Children and Teachers: Inspecting Knowledge Level of Individuals and Peers, in J.C. Lester, R.M. Vicari & F. Paraguacu (eds), *Intelligent Tutoring Systems: 7th International Conference*, Springer-Verlag, Berlin Heidelberg, pp. 646 - 655.

Bull, S. and Kay, J. (2005). A Framework for Designing and Analysing Open Learner Modelling. In, Proceedings of Learner Modelling for Reflection Workshop, held at the 12th International Conference on Artificial Intelligence in Education (AIED2005), Amsterdam 18-22 July, 2005, pp. 81-90.

Bull, S., Mangat, M., Mabbott, A., Abu Issa, A., Marsh, J. (2005) Reactions to inspectable learner models: sever year olds to university students. In, Kay, J., Lum, A. and Zapata-Rivera, D. (Eds.), Proceedings of Learner Modelling for Reflection Workshop, held at the 12th International Conference on Artificial Intelligence in Education (AIED2005), Amsterdam 18-22 July, 2005, pp. 91-99.

Bull, S., Mangat, M., Mabbott, A., Abu Issa, A.S. & Marsh, J. (2005b). Reactions to Inspectable Learner Models: Seven Year Olds to University Students, *Proceedings of Workshop on Learner Modelling for Reflection*, Int. Conference on Artificial Intelligence in Education 2005, pp. 1-10.

Bull, S. and Kay, J. (2006) Student Models that Invite the Learner In: The SMILIJ Open Learner Modelling Framework. Technical Report, University of Sydney, Australia, TR 580.

Calvi L., Cristea, A. (2002). Towards Generic Adaptive Systems: Analysis of a Case Study. De Bra, P., Brusilovsky, P., Conejo R. (Eds.), Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Malaga, Spain, May 2002, Springer:LNCS 2347, pp. 79-89.

Carroll, J.M. and Rosson, M.B. (1987). The paradox of the active user. In J.M. Carroll (Ed.), *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. Cambridge, MA: MIT Press. pp. 80-111.

Cheverst, K., Byun, H., Fitton, D., Sas, C., Kray, C., Villar, N. (2005) Exploring Issues of User Model Transparency and Proactive Behaviour in an Of .ce Environment Control System. User Modeling and User-Adapted Interaction (2005) v. 15, Springer, pp. 235 –273.

Cohen, P. R. and Perrault, C. R.: 1979, Elements of a Plan-based theory of speech acts. Cognitive Science 3, pp. 177-212.

Corbett, A.T., Anderson, J.R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4, pp. 253-278.

Cristea A., Aroyo, L. (2002). Adaptive Authoring of Adaptive Educational Hypermedia. De Bra, P., Brusilovsky, P., Conejo R. (Eds.), Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Malaga, Spain, May 2002, Springer:LNCS 2347, pp. 122-132.

Czarkowski, M. (1998) ATML, Tutor and ADAPT: an adaptive hypertext teaching system. Technical Report, TR 519. University of Sydney Australia. ISBN 1 86487 044 3.

Czarkowski, M. and Kay, J. (2000) Bringing scrutability to adaptive hypertext teaching. In: G. Gauthier, C. Frasson and K. VanLehn (eds.) Intelligent Tutoring Systems. Lecture Notes in Computer Science, Vol. 1839, (Proceedings of 5th International Conference on Intelligent Tutoring Systems (ITS'2000), Montreal, Canada, June 21-23, 2000) Berlin: Springer-Verlag, pp. 421-432.

Czarkowski, M., Kay, J. (2001) Tutor: support for scrutably personalised documents. ADCS 2001 Proceedings of the Sixth Australasian Document Computing Symposium, Volume 1, pp.29-36, Vercoustre, AM & Kay, J, Basse. 2001. ISBN 1-86487-440-6.

Czarkowski, M. and Kay, J. (2002) A scrutable adaptive hypertext. In: P. De Bra, P. Brusilovsky and R. Conejo (eds.) Proceedings of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002) Proceedings, Málaga, Spain, May 29-31, 2002, pp. 384-387.

Czarkowski, M., Kay, J. (2003a). How to give the user a sense of control over the personalization of adaptive hypertext? Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, User Modeling 2003 Session, pp. 121-132.

Czarkowski, M., Kay, J. (2003b). Challenges of Scrutable Adaptivity. Poster in: AIED 2003: proceedings from the 11th International Conference in Artificial Intelligence in Education; Sydney, Australia, July 20-24 Ulrich Hoppe, Felisa Verdejo, Judy Kay. IOS Press. pp. 404 - 407.

Czarkowski, M., Kay, J. and Potts, S. (2005a) Web Framework for Scrutable Adaptation. In: Proceedings of Workshop on Learner Modelling for Reflection, to Support Learner Control, Metacognition and Improved Communication between Teachers and Learners at 12th International Conference on Artificial Intelligence in Education, AIED'2005, Amsterdam, July 19, 2005, IOS Press, pp. 11-18.

Czarkowski, M., Kay, J. and Potts, S. (2005b) Scrutability as a core interface. Poster in Proceedings of 12th International Conference on Artificial Intelligence in Education, AIED'2005, Amsterdam, July 19, 2005, IOS Press, pp. 783-786.

Czarkowski, M. (2005c) Evaluating Scrutable Adaptive Hypertext. In proceedings from the 10th International conference on User Modelling, Edinburgh, UK, 24th-29th July 2005. Available online at http://www.easy-hub.org/hub/workshops, accessed 17th February 2006.

Czarkowski, M., Kay, J. (2006) Giving learners a real sense of control over adaptivity, even if they are not quite ready for it yet, Chen S and G Magoulas (eds), *Advances in Web-based Education: Personalized Learning Environments*, IDEA Publishing, Chapter 5, pp. 93-105.

De Bra, P., Calvi, L., (1997) Creating Adaptive Hyperdocuments for and on the Web. Proceedings of the AACE WebNet'97 Conference, pp. 149-154, Toronto, 1997.

De Bra, P., Aerts, A., Smits, D., Stash, N.,(2002) AHA! Version 2.0, More Adaptation Flexibility for Authors. Proceedings of the AACE ELearn'2002 conference, October 2002, pp. 240-246.

De Bra, P., Brusilovsky, P., Conejo, R. (Eds.) (2002) Proceedings from Adaptive Hypermedia and Adaptive Web-Based Systems, second International Conference, AH2002, Malaga, Spain, May, 2002, Springer, LNCS 2347.

De Bra, P., Aerts, A., Rousseau, B. (2002b). Concept Relationship Types for AHA! 2.0. In Proceedings of the AACE ELearn'2002 conference, October 2002, pp. 1386-1389.

De Bra, P., Aqerts, A., Berden, B., de Lange, B., Rousseau B., Santic, T., Smits, D. and Stash, N. (2003). AHA! The Adaptive Hypermedia Architecture. Proceedings of the International Conference on Hypertext and Hypermedia 2003, Nottingham (UK), pp. 81-84.

De Bra, P., Aerts, A., Berden, B., De Lange, B., Rousseau, B., Santic, T., Smits, D., Stash, N., (2003a) AHA! The Adaptive Hypermedia Architecture. Proceedings of the ACM Hypertext Conference, Nottingham, UK, August 2003. Refer online to http://aha.win.tue.nl, accessed 17th February 2006.

De Bra, P., Aerts, A., Berden, B., De Lange, B., (2003b) Escape from the Tyranny of the Textbook: Adaptive Object Inclusion in AHA!. Proceedings of the AACE ELearn 2003 Conference, Phoenix, Arizona, November 2003, pp. 65-71. Refer online to http://aha.win.tue.nl, accessed 17th February 2006.

De Bra, P., Santic, T., Brusilovsky, P., (2003c) AHA! meets Interbook, and more... Proceedings of the AACE ELearn 2003 Conference, Phoenix, Arizona, November 2003, pp. 57-64.

De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer.

De Bra, P., Aroyo, L., Chepegin, V. (2004b) The Next Big Thing: Adaptive Web-Based Systems. Journal of Digital Information (online journal), Volume 5 Issue 1, Article No. 247, 2004-05-27.

Diaz, A., Gervas P. (2004) Adaptive User Modelling for Personalization of Web Contents. In proceedings: De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer, pp. 65 – 74.

Dimitrova, V. (2003). StyLE-OLM: Interactive Open Learner Modelling, *International Journal of Artificial Intelligence in Education* 13(1), pp. 35-78.

Dolog, P., Henze, N., Nejdl, W., Sintek, M. (2004) The Personal Reader: Personalizing and Enriching learning Resources Using Semantic Web Technologies. In proceedings: De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer, pp. 85 – 94.

European Parliament Privacy Directive (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. The Official Journal of the European Communities of 23 November 1995 No L. 281 p. 31. An unofficial text is available at http://www.cdt.org/privacy/eudirective/EU_Directive_.html accessed November 2005.

Farzan, R., Brusilovsky, P. (2005) Social navigation support through annotation-based group modelling. In proceedings, L. Ardissono, P. Brna, A. Mitrovic (Eds.), 10th International conference on User Modelling (UM2005), Edinburgh, UK, 24th-29th July 2005, LNAI 3588, Springer-Verlag, pp. 463-472.

Freyne, J., Smyth, B. (2004) An Experiment in Social Search. In proceedings: De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer, pp. 95 – 103.

Hansen, C. & McCalla, G. (2003). Active Open Learner Modelling, *Proceedings of Workshop on Learner Modelling for Reflection (Supplemental Proceedings vol 5)*, International Conference on Artificial Intelligence in Education 2003, pp. 248-257.

Henderson, A. and Kyng,M. (1995) From customisable systems to intelligent agents, In: R.M. Baeker, J. Grudin, W. Buxton and S. Greenberg (eds), Readings in Human-Computer Interaction, Morgan Kaufmann, pp. 783-792.

Hohl, H., Böcker, B., Gunzenhäuser, R. (1995) Hypadapter: An adaptive hypertext system for exploratory learning and programming. In. Adaptive Hypertext and Hypermedia, User Modeling and User-Adapted Interaction, Springer, Volume 6, Numbers 2-3, pp. 131 - 156.

Höök, K., Karlgren, J., Waern, A., Dahlbäck, N., Jansson, C.J., Karlgren, K., and Lemaire, B. (1996). A Glass Box Approach to Adaptive Hypermedia. Journal of User Modeling and User Adapted Interaction, special issue on Adaptive Hypermedia, 6(2-3), pp. 157-184.

Höök, K. (2000) Steps to take before Intelligent User Interfaces become real, Journal of Interacting with Computers, vol. 12, no. 4, pp. 409-426.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In M. G. Williams, M. W. Altom, K. Ehrlich, & W. Newman (Eds.), Human factors in computing systems: CHI '99 conference proceedings. New York: ACM. pp. 159-166.

Jameson, A., Paris, C. and Tasso, C. (Eds.) (1997) User Modeling, Proceedings of the Sixth International Conference UM97, New York, 1997. Springer-Verlag. Available online at http://w5.cs.uni-sb.de/UM97, accessed 17th February 2006.

Jameson, A. (1999) Article titled “What Can the Rest of Us Learn From Research on Adaptive Hypermedia - and Vice-Versa?”. Comments by Anthony Jameson on the book Adaptive Hypertext and Hypermedia, edited by Peter Brusilovsky, Alfred Kobsa, and Julita Vassileva (Dordrecht: Kluwer, 1998). Available online at http://w5.cs.uni-sb.de/~jameson/ahh/ahh-comments.html, accessed 14th January 2006.

Jameson, A., Schwarzkopf, E. (2002). Pros and Cons of Controllability: An Empirical Study, De Bra, P., P. Brusilovsky, R. Conejo (Eds.), Proceedings of AH'2002, Adaptive Hypermedia and Adaptive Web-Based Systems, Springer, pp. 193-202.

Jameson, A. (2003) Adaptive interfaces and agents. In Jacko, J.A., Sears, A., eds.: Human-Computer Interaction Handbook. Erlbaum, Mahwah, NJ (2003) pp.305-330.

Jameson, A. (2003b) Systems That Adapt to Their Users: An Integrative Overview.
Tutorial presented at UM2003, the Ninth International Conference on User Modelling UM2003, Johnstown USA, 2003. Available online at http://dfki.de/~jameson/pdf/um03-tutorial-jameson.pdf (accessed 28 February, 2006).

Jameson, A. (2005) *User Modeling Meets Usability Goals.* Invited talk at UM 2005, the Tenth International Conference on User Modeling. July 26th, 2005.

Johnson, C. (2004) Innovations In Online Retail. Article by Forrester Research, December 9, 2004. Accessed online from http://www.forrester.com.

Johnson, W.L., Vilhjalmsson, H. & Marsella, S. (2005). Serious Games for Language Learning: How Much Game, How Much AI?, in C-K. Looi, G. McCalla, B. Bredeweg & J. Breuker, *Artificial Intelligence in Education*, IOS Press, Amsterdam, pp. 306-313.

Kay, J. and Kummerfeld, B. (1994). An Individualised Course for the C Programming Language. In Proceedings of the Second International WWW Conference '94 Mosaic and the Web. URL: http://www.cs.su.oz.au/%7Ebob/kay-kummerfeld.html, also at: http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Educ/kummer feld/kummerfeld.html, accessed 17th February 2006.

Kay, J. (1999) *A scrutable user modelling shell for user-adapted interaction*. PhD thesis, University of Sydney.

Kay, J. (2001) Learner Control. User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), Kluwer Academic Publishers, pp. 111-127.

Kay, J., B. Kummerfeld, and P. Lauder (2002). Personis: A Server for User Models. Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2002 Proceedings, 2002: pp. 203-212.

Kay, J. & Uther, J. (2003) VIUMm, a web-based visualisation of large user models. In, proceedings of the ninth international conference on User Modelling (UM2003), Brusilovsky, P. et al. (Eds.), Johnstown, PA, June 2003, pp. 198-202.

Kay, J. & Lum, A. (2005). Exploiting Readily Available Web Data for Scrutable Student Models, in C K. Looi, G. McCalla, B. Bredeweg & J. Breuker, *Artificial Intelligence in Education*, IOS Press, Amsterdam, pp. 338-345.

Kay, J., Lum, A. and Zapata-Rivera, D. (Eds.) (2005). Proceedings of Learner Modelling for Reflection Workshop, held at the 12th International Conference on Artificial Intelligence in Education (AIED2005), Amsterdam 18-22 July, 2005.

Kay, J. and Niu, W. (2005) Adapting Information Delivery to Groups of People. In proceedings from workshop on New Technologies for Personalized Information Access, (PIA 1005), Brusilovsky, P., Callaway, C., Nurnberger, A. (Eds.), held in conjunction with the tenth Internalional Conference on User Modeling (UM2005) Edinburgh, 24th-29th July, pp.34-43.

Kobsa, A. & Josef, F. (2000) A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web, User Modeling and User-Adapted Interaction 10. Kluwer Academic Publishers, pp. 209-249.

Kobsa, A. (2001) Generic User Modeling Systems. User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), Kluwer Academic Publishers, pp. 49-63.

Kobsa, A. (2002) Personalized Hypermedia and International Privacy. Communications of the ACM 45(5), Special Issue on Adaptive Web-based Systems and Adaptive Hypermedia, May 2002, pp. 64-67.

Kobsa, A., Cranor, L. (Eds.) (2005) Workshop on Privacy Enhanced Personalization (PEP2005), held at UM2005, 10th International conference on User Modelling, Edinburgh, UK, 24th-29th July 2005.

Koch, N., Wirsing, M. (2002) The Munich Reference Model for Adaptive Hypermedia Applications. In proceedings: De Bra, P., Brusilovsky, P., Conejo, R. (Eds.) (2002) Proceedings from Adaptive Hypermedia and Adaptive Web-Based Systems, second International Conference, AH2002, Malaga, Spain, May, 2002, Springer, LNCS 2347, pp. 213 – 222.

Kummerfeld, B. and P. Lauder (2003). "Cellerator", http://www.it.usyd.edu.au/~bob, accessed August 2003.

Li, L. & Kay, J. (2005). Assess: Promoting Learner Reflection in Student Self-Assessment, *Proceedings of Workshop on Learner Modelling for Reflection*, International Conference on Artificial Intelligence in Education 2005, pp. 32-41.

Maes, P., Schneiderman, B. (1997). Direct Manipulation vs. Interface Agents: A Debate. Interactions, Volume IV Number 6, November 1997, ACM Press, pp. 42-61.

Manber, U., Patel, A. & Robinson, J. (2000). Experience with personalization on Yahoo!. Communications of the ACM, Aug 2000; Vol. 43(8), pp. 35-39.

Mazza, R. & Dimitrova, V. (2003). CourseVis: Externalising Student Information to Facilitate Instructors in Distance Learning, in U. Hoppe, F. Verdejo & J. Kay (eds), *Artificial Intelligence in Education*, IOS Press, Amsterdam, pp. 279-286.

Mitrovic, A. & Martin, B. (2002). Evaluating the Effects of Open Student Models on Learning, in P. De Bra, P. Brusilovsky & R. Conejo (eds), *Adaptive Hypermedia and Adaptive Web-Based Systems, Proceedings of Second International Conference*, Springer-Verlag, Berlin Heidelberg, pp.296-305.

Morales, R., Pain, H., Bull, S. and Kay, J. (Eds.) (1999). Open, interactive, and other overt approaches to learner modeling. Proceedings of the workshop at AIED 1999. International Journal of Artificial Intelligence in Education, 10, pp. 1070-1079.

Nelson, T. H. (1965) A File Structure for the Complex, The Changing and The Indeterminate, ACM 20th National Conference, 1965.

Nielsen, J. (1990) Hypertext and Hypermedia. Published by Academic Press, Boston, MA, 1990. ISBN 0-12-518410-7.

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. International Journal of Human-Computer Studies, 41(1-6), pp. 385-397.

Nielsen, J. (1994b). Usability Engineering. Published by Morgan Kaufmann, San Francisco, 1994.

Nunes, José Miguel Baptista & Fowell, Susan P. (1996) "Hypermedia as an experiential learning tool: a theoretical model". Information Research, 2(1) Available at: http://InformationR.net/ir/2-1/paper12.html, accessed 24 February 2006.

Paramythis, A., Totter, A., and Stephanidis, C. (2001). A Modular Approach to the Evaluation of Adaptive User Interfaces. In Stephan Weibelzahl, David Chin, and Gehard Weber (Eds.) *Empirical Evaluation of Adaptive Systems.* Proceedings of workshop held at the Eighth International Conference on User Modeling in Sonthofen, Germany, July 13th, 2001, pp. 9-24.

Perrault, C. R., Allen, J. F. and Cohen, P. R.: 1978, Speech acts as a basis for understanding dialogue coherence. Report 78-5, Department of Computer Science, University of Toronto, Canada.

Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H. & Magoulas, G.D. (2003). Personalizing the Interaction in a Web-Based Educational Hypermedia System: The Case of INSPIRE, *User-Modeling and User-Adapted Interaction* 13(3), pp. 213-267.

Peng, X., Silver, D. (2005) User Control over User Adaptation: A Case Study. In proceedings, L. Ardissono, P. Brna, A. Mitrovic (Eds.), 10th International conference on User Modelling (UM2005), Edinburgh, UK, 24th-29th July 2005, LNAI 3588, Springer-Verlag, pp. 443-487.

Personalisation Survey (2000), accessed online http://www.personalization.org/SurveyResults.pdf on 28th October 2005.

Rich, E. (1979a), Building and Exploiting User Models. Ph.D. Thesis, Department of Computer

Science, Carnegie-Mellon University, Pittsburgh, PA.

Rich, E, (1979b). User modeling via stereotypes. Cognitive Science, 3, pp. 355-66.

Self, J. (1999). Open Sesame?: Fifteen variations on the theme of openness in learning environments, International Journal of Artificial Intelligence in Education 10, pp. 1020-1029.

Schwab, I. and Kobsa, A. (2002). Adaptivity through Unobstrusive Learning. KI 3(2002), Special Issue on Adaptivity and User Modeling, pp. 5-9.

Tanimoto, S. (2005). Dimensions of Transparency in Open Learner Models. In, Proceedings of Learner Modelling for Reflection Workshop, held at the 12th International Conference on Artificial Intelligence in Education (AIED2005), Amsterdam 18-22 July, 2005, pp. 100-106.

Tongchai, N., Brna, P. (2005) Enhancing metacognitive skills through the use of a group open learner model based on the zone of proximal development. In, Kay, J., Lum, A. and Zapata-Rivera, D. (Eds.), Proceedings of Learner Modelling for Reflection Workshop, held at the 12th International Conference on Artificial Intelligence in Education (AIED2005), Amsterdam 18-22 July, 2005, pp. 91-99.

Tsandilas, T., Schraefel, M. C. (2004) Usable adaptive hypermedia systems. New Review of Hypermedia and Multimedia, Vol. 10, No. 1, June 2004, pp. 5-29.

Vannevar, B. (1945) “As we may think”. Article in The Atlantic Monthly, July 1945 issue. Reproduced online at http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush.shtml, accessed 17th February 2006.

Vrieze, P., van Bommel, P., van der Weide, T. (2004) A Generic Adaptivity Model in Adaptive Hypermedia. . In proceedings: De Bra, P.,Nejdl, W. (Eds.) (2004) Adaptive Hypermedia and Adaptive Web-Based Systems. Proceedings from Third International Conference, AH2004, Eindhoven, The Netherlands, August 23-26, 2004. Lecture Notes In Computer Science, 3137, Springer, pp. 344 – 347.

Wahlster, W.and Kobsa, A.(1986) Dialogue-based user models. Proceedings of the IEEE, 74(7), pp. 948–960.

Weber, G., Brusilovsky, P., Schwarz, O. (1996). ELM-ART: an intelligent tutoring system on the World Wide Web. Frasson, C., Gauthier, G., Lesgold, A. (Eds.), Proceedings of the Third International Conference on Intelligent Tutoring Systems, ITS-96. Springer, Berlin, pp. 261-269.

Weber, G., Brusilovsky, P. (2001) ELM-ART: An adaptive versatile system for Web-based instruction. *International Journal of Artificial Intelligence in Education* 12 (4), Special Issue on Adaptive and Intelligent Web-based Educational Systems, pp. 351-384.

Weibelzahl, S., Paramythis, A., Masthoff, J. (Eds.) (2005) Workshop on Evaluation of Adaptive Systems, held at UM2005, 10th International conference on User Modelling, Edinburgh, UK, 24th-29th July 2005. Available online at http://www.easy-hub.org/hub/workshops, accessed 25 February 2006.

Zapata-Rivera, J-D. & Greer, J.E. (2001). Externalising Learner Modelling Representations, *Proceedings of Workshop on External Representations of AIED,* International Conference on Artificial Intelligence in Education 2001, pp. 71-76.

Zapata-Rivera, J-D. & Greer, J.E. (2004). Interacting with Inspectable Bayesian Student Models. International Journal of Artificial Intelligence in Education 14 (2004), pp 127-163.

Zapata-Rivera, J-D., Underwood, J.S. & Bauer, M. (2005). Advanced Reporting Systems in Assessment Environments, *Proceedings of Workshop on Learner Modelling for Reflection*, International Conference on Artificial Intelligence in Education 2005, pp. 23-31.

Zaslow, Jeffrey (2002). “Oh no! My TiVo thinks I'm gay”, The Wall Street Journal, Dec. 04, 2002.