Journal of Universal Computer Science, vol. 4, no. 4 (1998), 429-448
submitted: 10/12/97, accepted: 28/12/97, appeared: 28/4/98 © Springer Pub. Co.

A Study of User Model Based Link Annotation in Educational Hypermedia

Peter Brusilovsky
(Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh PA 15213, USA
plb@cs.cmu.edu)

John Eklund
(Learning Systems Research and Development Group, Faculty of Education, The University of Technology, Sydney NSW Australia
j.eklund@uts.edu.au)

Abstract: Adaptive link annotation is a new direction within the field of user-model based interfaces. It is a specific technique in Adaptive Navigation Support (ANS) whose aim is to help users find an appropriate path in a learning and information space by adapting link presentation to the goals, knowledge, and other characteristics of an individual user. More specifically, ANS has been implemented on the WWW in the *InterBook* system as link annotation indicating several states such as visited, ready to be learned, or not ready to be learned. These states represent an expert's suggested path for an individual user through a learning space according to both a history-based (tracking where the user has been), and a pre-requisite based (indexing of content as a set of domain model concepts) annotation. This particular process has been more fully described elsewhere [Brusilovsky, Eklund & Schwarz 1998].

This paper details results from an investigation to determine the effectiveness of user-model based link annotation, in a real-world teaching and learning context, on learning outcomes for a group of twenty-five second year education students in their study of databases and spreadsheets. Using sections of a textbook on *ClarisWorks* databases and spreadsheets, which had been authored into the *InterBook* program, students received sections of the text both with and without the adaptive link annotation. Through the use of audit trails, questionnaires and test results, we show that while this particular form of ANS implemented in *InterBook* initially had a negative effect on learning of the group, it appears to have been beneficial to the learning of those particular students who tended to accept the navigation advice, particularly initially when they were unfamiliar with a complex interface. We also show that ANS provided learners with the confidence to adopt less sequential paths through the learning space. Considering ANS tools comprised a minimal part of the interface in the experiment, we show that they functioned reliably well. Discussion and suggestions for further research are provided.

Keywords: Hypertext, Adaptive Hypermedia, Navigation support, Evaluation, Navigation, User Model, WWW

1 Introduction

Adaptivity is one of the ways of increasing the functionality of hypermedia. Adaptive hypermedia systems build a model of the goals, preferences, and knowledge of the individual user and use this throughout the interaction for accommodating the individual needs of the particular user. Adaptive hypermedia can be useful in any situation when the system is expected to be used by people with different goals and knowledge, where the hyperspace is reasonably big, or where the system can successfully guide the user in his or her work [Brusilovsky 1996]. Education is one of the most promising application areas for adaptive hypermedia, as it can be applied to adapt the presented information to the current knowledge level of the student, to provide navigation support, and to guide the student in the learning process without being too prescriptive and directive.

There are two general methods of implementing adaptation in adaptive hypermedia: *adaptive presentation* (or content-level adaptation) and *adaptive navigation support* (or link-level adaptation). In adaptive presentation the content of a hypermedia page is generated or assembled from pieces according to the user's background and knowledge state. Generally, qualified users receive more detailed and deep information, while novices receive more additional explanation. By adaptive navigation support (ANS) we mean all the methods of altering visible links to support hyperspace navigation.

Adaptive annotation of links is a promising technique for ANS ineducational hypermedia. This technique was suggested in [Brusilovsky, Pesin & Zyryanov 1993; de La Passardiere & Dufresne 1992]. The idea of *adaptive annotation* technology is to augment the links with some form of comments which can tell the user more about the current state of the nodes behind the annotated links. These annotations can be provided in textual form [Zhao, O'Shea & Fung 1993] or in the form of visual cues using, for example, different icons [Brusilovsky, Schwarz & Weber 1996; de La Passardiere & Dufresne 1992], colours [Brusilovsky & Pesin 1994], font sizes [Hohl, Bšcker & GunzenhŠuser 1996], or font types [Brusilovsky, Schwarz & Weber 1996]. Annotation seems to be a very relevant way of adaptive navigation support in educational hypermedia. Annotation can be naturally used with all possible forms of links in hypertext and hypermedia. This technique supports stable order of links and avoids problems with incorrect mental maps.

Our position is that adaptive navigation support can be successfully applied in educational hypermedia in a real world teaching and learning context and that adaptive annotation is a relevant technique for that purpose. However at present there are very few instructional systems with adaptive navigation support and there are very few experimental studies which can test how useful adaptive navigation support can be for educational application. [Weber & Specht 1997] used the ELM-ART [Brusilovsky, Schwarz & Weber 1996] system to count the number of navigation steps for those with and without ANS and found no significant difference with a relatively small (n=16) group of novice learners. [Brusilovsky & Pesin 1998] investigated ISIS-Tutor system [Brusilovsky & Pesin 1994] and reported a significant decrease of the number of navigation steps and the number of repeated visits to the same node for a group with ANS. We focus instead on the impact of ANS on learning outcomes and user paths. In the following sections we describe the *InterBook* system which demonstrates a particular implementation of adaptive navigation support, and we report the results of an empirical study using *InterBook*.

2 Adaptive Navigation Supportin *InterBook*

InterBook is a system for authoring and delivering adaptive electronic textbooks on the WWW. Electronic textbooks reside on an InterBook server and can be accessed with any frame-enabled Web browser. The *InterBook* interface [Fig. 1] divides the screen into four sections, the largest window being *the textbook window* in which the content in the form of text, hypertext and graphics appears. On the top-right is the toolbar, in which the links to a table of content,a glossary, and a search interface, and a help button appear. The window at the top left is called the navigation bar, and this provides the learner with a navigable hierarchy of surrounding nodes. The window at the bottom right is called the concept bar, and this lists the *pre-requisite and outcome concepts* for the section presented in the textbook window.

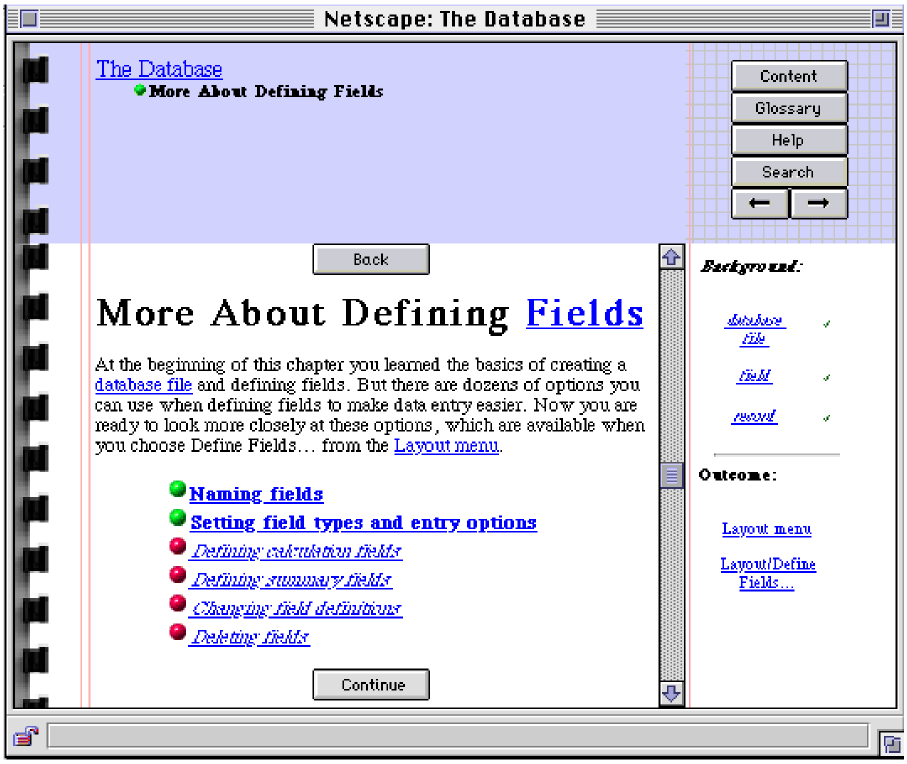


Figure 1: Text window of InterBook with adaptive link annotation. Green bullet means recommended, red bullet means "not ready to be learned", white bullet means "nothing new", while a checked bullet means "visited".

The key to adaptivity in InterBook is what we call "knowledge behind pages". InterBook uses a structured *domain model* represented as a network of *domain concepts*. Domain concepts are important terms of the domain. They designate atomic pieces of knowledge about the domain. A special part of an electronic textbook, the Glossary provides some descriptions of domain model concepts.A description of each concept is individually accessible as a glossary entry [Fig. 2].

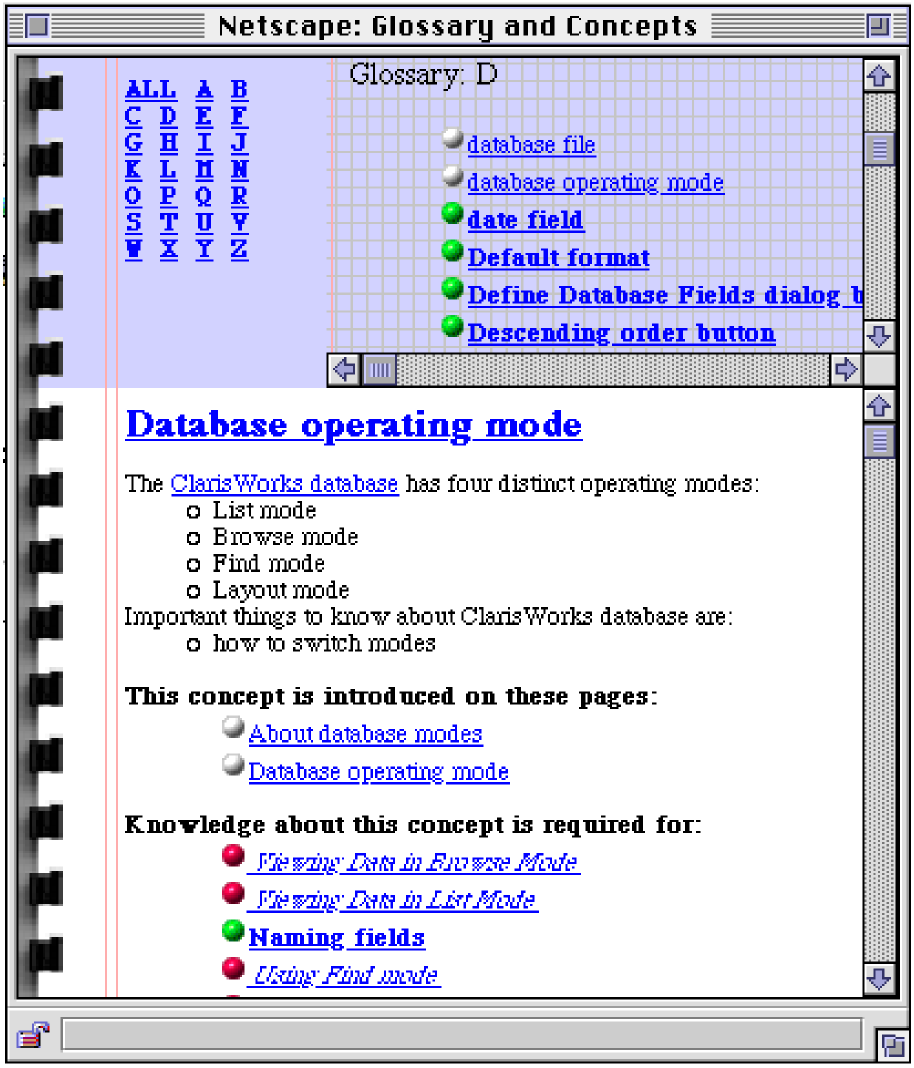


Figure 2: Glossary window of InterBook showing a glossary entry for the concept "Database operating mode"

All sections of an electronic textbook are indexed with domain model concepts. For each section, a list of concepts related with this section is provided (this list is called the spectrum of the section). The spectrum can also represent the role of a concept in the section (either an outcome concept or a background concept). A concept is included in the spectrum as an outcome concept if some part of this section presents the piece of knowledge designated by the concept. A concept is included in the spectrum as a prerequisite concept if a student has to know this concept to understand the content of the section.

The knowledge about the domain and about the textbook content is used by *InterBook* to serve a well-structured hyperspace. In particular, *InterBook* generates links between the glossary and the textbook. Links are provided from each textbook section to corresponding glossary entries for each involved background or outcome concept. Similarly, for each glossary entry describing a concept *InterBook* provides links to all textbook units that can be used to learn this concept. This means that an *InterBook* glossary integrates features of an index and a glossary.

InterBook uses coloured bullets and different fonts to provide adaptive navigation support [Fig. 1]. Wherever a link appears on *InterBook* pages (in the table of contents, in the glossary or on a regular page), its font and the colour of its bullet will inform the user about the status of the node behind that link. Green bullet and bold font means "ready and recommended", i.e., the node is ready-to-be-learned but still not learned and contains some new material. A red bullet and an italic font warn about a not-ready-to-be-learned node. A white bullet means "clear, nothing new" (i.e., all concepts presented on a node are known to the user). A check mark is added for already visited nodes. *InterBook* integrates all three methods of annotation: history-based (on the basis of where the user has been), prerequisite-based (on the basis of what prerequisite nodes the user has visited, and knowledge-based (on the basis of the user's demonstrated knowledge).

The user model in *InterBook* represents levels of user's knowledge of every domain concept. It is initialized from the registration page via a stereotype model, and is modified as the user moves through the information space. The user model for each user is stored in a file on the server.

3 Experimental Overview

In a study involving 25 undergraduate teacher education students in an educational computing elective at the University of Technology, Sydney, students were exposed to two chapters of a textbook [Rubin 1996] about *ClarisWorks* databases and spreadsheets, and used the *InterBook* system both *with* [Fig. 1] and *without* [Fig. 3] adaptive link annotation. The experiment was created to be in a real-world teaching and learning context, with the use of *InterBook* as an integral part of a university subject as described inthe previous section.

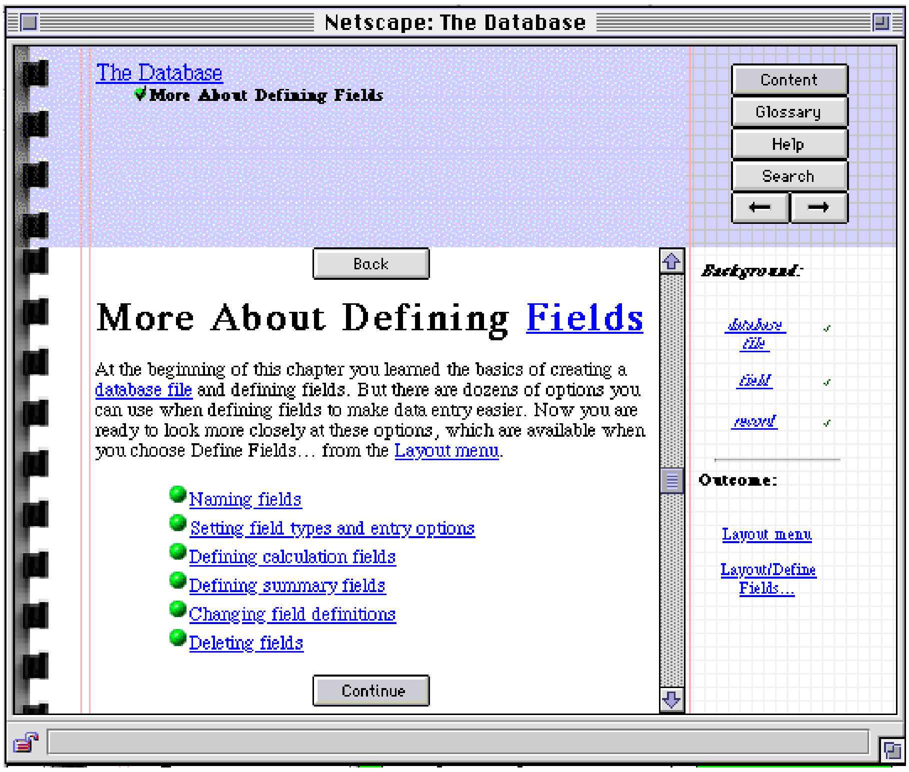


Figure 3: Text window of the non-adaptive version of InterBook used for the experiment. Adaptively chosen bullets and fonts are replaced by green bullets and regular font.

The goal of this experiment was to assess what impact, if any, user-model based link annotation would have on students' learning and on their paths through the learning space, in this realistic situation. The experiment was aimed to investigate both the effect of link annotation on learning and the effect of link annotation on user paths. The hypothesis was that adaptive link annotation would provide students with a more efficient path through the knowledge space with improved learning outcomes. Tests of knowledge were carried out, audit trails and questionnaires were gathered and the results analyzed.

The experiment took place over a four-week period. In the first two-hour session, students were introduced to *InterBook* and its features explained to them. They used the system for an hour, and answered a questionnaire about its features. This questionnaire showed that almost all students were familiar with what each of the buttons and annotations meant. They were then free to use the system at any time during the following week.

In the second session, students were randomly divided into two groups of equal size, one group receiving the link annotation, while the other group did not [Tab. 1]. They were allowed access to the chapter of the textbook on databases which had been authored into *InterBook* and they completed a questionnaire. Students had access to the database chapter for the following week.

Group 1	**Group 2**
Database chapter WITH adaptive link annotation n= 12	Database chapter WITHOUT adaptive link annotationn=13
Spreadsheet chapter WITHOUT adaptive link annotation n=12	Spreadsheet chapter WITH adaptive link annotation n=13

Table 1: Allocation of adaptive link annotation to groups

In the third session, students took a multiple choice test on the database section of the textbook. They were then allowed access to the spreadsheet section of the textbook in *InterBook* which they could access for the following week. This time, the group that did have the adaptive link annotation for the database section now did not receive it, and vice-versa for the other group [see Tab. 1]. In the final session, students took a multiple choice test on the spreadsheet section and completed a questionnaire. The audit trails from the sessions were extracted, and analyzed along with the test results and the questionnaire responses. After the second session the students were asked to rate their use of the various features of the interface, apart from the link annotation which some of them had not been receiving. These ratings are shown in [Tab. 2]. The purpose of this was to determine if any feature was not well received by the students.

4 Experimental Results -Interface

A questionnaire was used to assess the functionality of each of the key interface features of *InterBook*, the results showing that all the features were working as expected, quite uniformly across the group.

Question	Level of response	Mean	Standard Deviation
drth The Multiple windows were	1=useless 5=useful	3.5	1.1
I used the hot links in the Text	1=never 5=often	3.6	1.0
The ’you are/were here feature’ in the table of content was	1=useless 5=useful	3.9	1.2
The Navigation was rather	1=hard 5=easy	3.3	1.2
I used the search feature	1=never 5=often	3.4	1.2
I took into account the checked balls	1=never 5=often	4.0	1.1
The search feature was	1=useless 5=useful	3.5	1.1
The local overview in the table of contents was:	1=useless 5=useful	3.2	0.9
The list of related pages in glossary was	1=useless 5=useful	4.1	1.0

Table 2: User’s rating of the interface features of InterBook (n=25)

5 Experimental results - Test Scores

5.1 Procedure

The students’ test scores were then used as a measure of their learning of the material in each of the sessions. There was a reasonable margin for error in using this variable, as students’ prior knowledge of the domain, and their learning of it from other sources such as the actual text held in the library’s closed reserve, could not be determined. It was particularly important that students’ scores in each test were a reliable measure of their learning, considering each test was rather short and they were not standard experimental instruments. Experimental effects which were not totally accounted for were minimized through two methods. Firstly the tests, consisting of ten and twelve multi-choice questions taken directly from the material in *InterBook*, were validated on another small group of fifteen students. A Cronbach Alpha for the database test was calculated at 0.58 and for the spreadsheet test an initial Cronbach Alphavalue of 0.23 was obtained, and these unsatisfactory results were improved by modifying each test. Adequate reliability and performance of all test items were established by discarding some of the test questions or individual distracters. In this way, each of the test questions was constructed to be an adequate predictor of how a student would score in the overall test, as is desirable in a norm-referenced test. At the conclusion of the experiment, alphas of 0.75 and 0.82 for thirty-two students were obtained for the database and spreadsheet tests respectively, and these very acceptable values were interpreted as establishing adequate reliability for the tests to be used in this and any subsequent experiment.

Secondly, other measures of performance were examined for the groups. To avoid the possibility that the randomly chosen groups consisted of a disproportionate number of less able or more able students, means and standard deviations of test results for each group (i.e., the group that was to receive the annotations and the group that was not), taken from other aspects of the course were compared, and these were found to be very similar. In other words the random selection of students provided two groups with very similar academic ability.

5.2 Results

The results of the students' knowledge tests are shown in table 3:

Group	**Test Result Database**	**Test Result Spreadsheet**
1 ANS on database only	6.41	7.77
2 ANS on spreadsheet only	7.12	8.10

Table 3: Test results for groups with and without ANS

A two-sample t-Test was performed on the results. The t value of -0.3667 shows that link annotation had a statistically significant *negative* effect at the $p<0.05$ level on the database session (the first session), and no effect on the spreadsheet session (the second session). This unexpected initial result suggested that further investigation was required.

A careful analysis of the audit trails revealed two factors. First, some of the subjects apparently learned about ClarisWorks from other sources since they were able to obtain good test results after hitting only 5 to 15 sections of the electronic textbook (the Databasepart alone contains about 100 sections). Second, for most of the students ANS appears to be a minor factor because about 80% of all navigation steps were made with Continue and Back buttons which were *not* annotated in the experimental version of *InterBook* [see Chapter 6 for more details]. In this situation we had touse some more elaborate techniques to find a relationship between ANS and test performance.

To exclude students who learned less from the system than from other sources in the study, two subgroups were introduced, the first based on spending a 'reasonable time' with the system. This consisted of those students who spent a reasonable time using *InterBook* over both sessions, as it became clear from the audit trails that a number of students relied heavily on either their previous knowledge of the content, or on the printed version of the *ClarisWorks* book. For both the database and spreadsheet sections, two-sample t-Tests showed that there was no significant difference at the 0.05 level in the test means for those with ANS and those without ANS.

The second (overlapping) subgroup was intended to eliminate those students who made very few (fewer than 15) hits on the system.

	Mean	Std. Dev.	Std. Error	Count	Minimum	Maximum	# Missing
Score, Total	7.9	1.2	.3	17	6.0	10.0	0
Score, yes	7.9	1.0	.4	8	7.0	9.0	0
Score, no	8.0	1.5	.5	9	6.0	10.0	0

Table 4: Database Subgroup which excludes those with fewer than 15 hits on the system

Table 4 shows that of the 17 students who made greater than 15 hits in the database section, 8 students received ANS and 9 students did not. There is no statistically significant difference in the test performance of those that did and those that did not receive ANS.

This result is natural taking into account that the average number of navigation steps (with annotated links) made by the "adaptive" group was too small to affect their performance. However, users appeared to be very different in their navigation behavior. Some of them almost never used annotated links, some of them used it reasonably often. We decided to investigate the performance of users who did use annotated links.

5.3 The Value of ANS for Those Who Use It

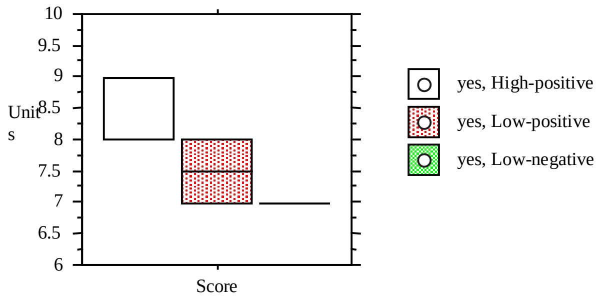
Separate audit trails for each of the two time periods were generated, to examine how users navigated through *InterBook* with and without ANS. For each user these trails showed the number of times they selected a link with a green ball and also a red ball, as well as their use of all the other features of the interface. Certain unexpected behaviour was immediately apparent for a small group of students, who were purposefully and continually selecting nodes which were not recommended. More generally, it was noted that just because link annotation was evident in the interface for one group, individual students within that group were accepting it to varying extents. Just because a student was offered link annotation does not mean that they were accepting or making profitable use of it. A measure of the students' acceptance of navigational advice was calculated from the audit trails, taken as the number of green ball hitsminus the number of red ball hits divided by the total hits. This measure of acceptance of this particular interface functionality is a more important variable than the fact that they were provided with it.

agreement rate = $(N_{greenballs} - N_{redballs})/(N_{greenballs} + N_{redballs} + N_{whiteballs})$

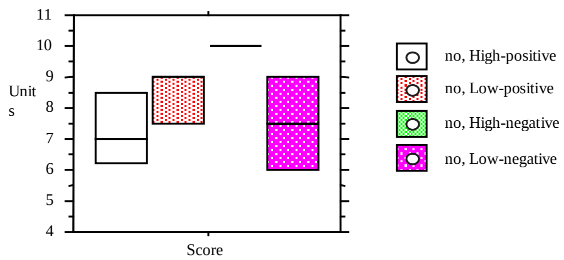
Where $N_{greenballs}$, $N_{redballs}$, and $N_{whiteballs}$ is the number of times the user hits a link which should be annotated with green, red, or white bullet. Note that for the students of a group with ANS these links were *really* annotated with bullets of different colours, while the students of anon-ANS group were shown the same green bullet regardless of the state.

For students who always follow green balls the agreement rate is 1, for those who always follow red balls it is -1. Then four distinct groups for each of with and without ANS were established depending on this agreement rate:

High-positive	rate > 0.5
Low positive	0 < rate <=0.5
Low negative	-0.5 < rate <= 0
High-negative	rate <= -0.5



(a)



(b)

Figure 4: Clustering acceptance of link annotation and database scores for the group with (a) and without (b) ANS. For the ANS group a better agreement rate results in generally better test results. Columns and keys have the same order, i.e., the top key corresponds to the most left column and the bottom key corresponds to the most right column.

A clear correlation (R=0.670) was found between the agreement rate and score in the database tests: the more students agree with system's suggestion, the better is the score - for the group receiving link annotation [Fig. 4a]. Moreover, in the group with link annotation there were no high-negative students at all. For the non-ANS group [Fig. 4b] there is a mild negative correlation (R=-0.383) and one high-negative student. This is natural because they have not seen the annotations, but it is also an argument for ANS - without annotation the students cannot recognize and use the state of the page.

The above calculations were repeated with a modified formula, namely:

agreement rate = ($N_{greenballs}$ - $N_{redballs}$)/All-Hits

This formula may provide a more reliable measure of agreement for the users who almost never hit annotated links (i.e., when $N_{greenballs}$ + $N_{redballs}$ + $N_{whiteballs}$ is very small). A similar correlation of 0.618 was obtained for the group with the ANS and -0.176 for those without ANS.

This positive correlation in the first session on databases suggests that while link annotation is a distracting complication to an interface, it is helpful to those that choose to follow it in terms of improving their knowledge of the content.

6 Student's Use of Individual Navigation Tools

One of the major problems in determining if ANS was effective or not was that approximately 80% of the available navigation tools that were used in the experimental version of *InterBook* were non-adaptable. An analysis of the proportion of use of different navigation tools [Tab. 6] shows that the non-annotated Continue button (pCONTINUE) is used more than all other navigation tools combined. The bullet-annotated pCONTENT, pINTRODUCING, pREQUIRING, pSEARCH, pREGISTER,pPATH, pHELP [see Tab. 5] are used on less than 20% of hits (for those users who received annotation). Additionally the checkmark-annotated pPREREC and pOUTCOME are almost not used at all with only 1-2% of hits. This implies that the majority of navigation choices made by students were made without annotations.

	Mean	Std. Dev.	Std. Error	Count	Minimum	Maximum	# Missing
pCONTINUE	15.4	8.6	2.1	17	2.0	30.0	0
pCONTENT	3.9	3.8	.9	17	1.0	16.0	0
pTEXT	2.9	3.1	.8	17	0.0	11.0	0
pBOOKSEL	1.5	.7	.2	17	1.0	3.0	0
pINTRODUCING	1.1	1.4	.3	17	0.0	4.0	0
pBOOKTITILE	.8	.8	.2	17	0.0	2.0	0
pTOOLBAR	.4	.9	.2	17	0.0	3.0	0
pOUTCOME	.4	1.0	.2	17	0.0	4.0	0
pREQUIRING	.3	.7	.2	17	0.0	2.0	0
pHELP	.1	.2	.1	17	0.0	1.0	0
pPATH	.1	.2	.1	17	0.0	1.0	0
pPREREQ	.1	.2	.1	17	0.0	1.0	0
Not annotated	21.9	7.8	1.9	17	8.0	40.0	0
Balls	5.6	4.4	1.1	17	1.0	19.0	0
Checkmarks	.4	1.1	.3	17	0.0	4.0	0

Table 5: The number of times different navigation tools were used [see Tab. 6]

Not annotated in both versions	**Annotated in ANS version**
- pCONTINUE - using continue button	- pCONTENT - link from the separate or embedded table of contents
- pBACK - using back button	-pINTRODUCING - link from a glossary page to a page introducing a concept
- pTEXT - hypertext reference from one page to another	- pREQUIRING -link from a glossary page to a page which requires a concept
- pBOOKSEL - link to a book from book list (top of the table of content page)	-pHELP - link to one of the helpful pages from background help page
- pBOOKTITLE - using the link to the book title in navigation centre	-pPATH - link to an higher level section from the navigation center
- pTOOLBAR - using buttons on the Toolbar	- pPREREC, pOUTCOME - links from the concept bar to glossary annotated with checkmarks

Table 6: Description of most often used navigation tools

Using sequential navigation (i.e., continue-back) vs. non-sequential navigation is known behaviour exposed by novices in hyperspaces. However, we were able to show that ANS encourages the novices to use annotated non-sequential tools more often. This was achieved using the count of hits on annotated links such as table of contents links versus non-annotated links such as Continue button.

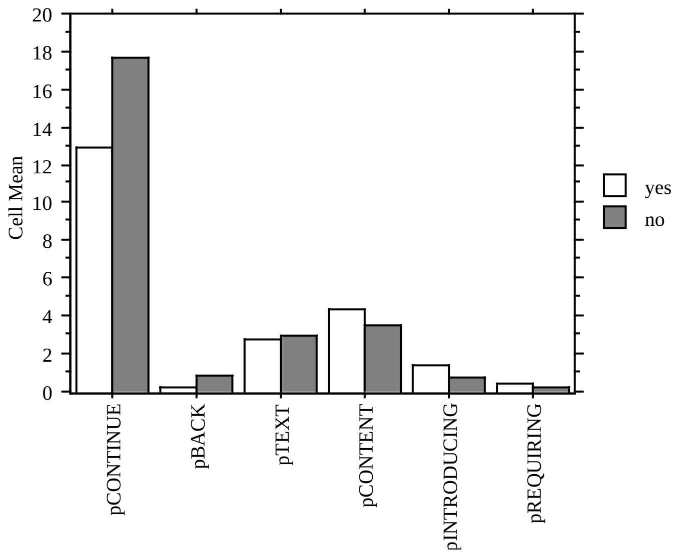


Figure 5: ANS and the use of sequential and non-sequential navigation tools (average number of hits per user). The non-shaded "yes" columns are for the ANS group; the shaded "no" columns are for the non-ANS group

[Fig. 5] and [Fig. 6] show that those students who did not receive ANS (the shaded "no" columns) used more of the non-annotatable and sequential navigation features (pCONTINUE, pBACK). At the same time those who did receive ANS (the non-shaded "yes" column) used more of the annotatable navigation features (pCONTENT, pINTRODUCING, pREQUIRING). Even the use of pTEXT which is a non-annotatable but non-sequential navigation tool is slightly smaller for the ANS group. This implies that ANS provides the learner with a non-linear guide through the learning space, and learners are more likely to use non-sequential paths with adaptive link annotation. It again reflects the student's trust in the annotations - ANS provides some security for those users who would like to follow non-linear paths but might be afraid of becoming lost.

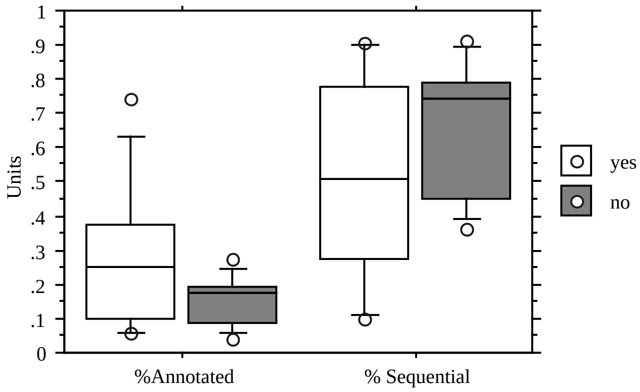


Figure 6: ANS and the proportion of use of annotatable (left) and sequential (right) navigation tools. The presence of ANS encourages the subjects of the ANS group (white "yes" bars) to use annotatable and non-sequential links more often.

7 The Role of The Page State

If the number of hits on pages of various 'states' [Tab. 7] is examined, it is clear that students prefer to visit ready-to-be-learned pages (s2) than those which are annotated as "no information" (s1). Students spent approximately twice as much time reading ready-to-be-learned pages than reading all other pages (s1 and s3) combined. Thus, the data shows that a green and unchecked page is one that students read most. This is naturally due to the fact that the Continue button was used most of the time, bringing the user into a page with a "ready to be learned" status.

	Mean	Std. Dev.	Std. Error	Count	Minimum	Maximum	# Missing
s1	4.294	2.687	.652	17	1.000	9.000	0
s2	15.882	6.343	1.538	17	6.000	30.000	0
s3	7.706	4.413	1.070	17	1.000	15.000	0
s1t	207.882	178.126	43.202	17	23.000	592.000	0
s2t	1220.765	515.682	125.071	17	451.000	2242.000	0
s3t	749.941	427.570	103.701	17	25.000	1476.000	0
r0	24.176	8.118	1.969	17	15.000	45.000	0
r1	3.706	2.779	.674	17	0.000	9.000	0
r0t	1944.294	518.402	125.731	17	1108.000	2935.000	0
r1t	234.824	207.745	50.386	17	0.000	748.000	0

Table 7: Number of hits on pages of various states [see Tab. 8]

Time	**Hits**
s1t = time on ready but not suggested (while ball)	s1 = hits on pages ready but not suggested (while ball)
s2t = time on ready and suggested (green ball)	s2 = hits on pages ready and suggested (green ball)
s3t = time on not ready (red ball)	s3 = hits on pages not ready (red ball)
r1t = time on pages that have been read before (checkmark over ball)	r0 = hits on pages that have not been read before (no checkmark overball)
r0t = time on pages that have not been read before (no checkmark over ball)	r1 = hits on pages that have been read before (checkmark over ball)

Table 8: Key to states of pages

Moreover, it can be seen from [Fig. 7] showing the average time students spent on different types of pages, that "nothing new", "not ready" and "ready" pages are very different. The average time spent on a not-ready page is much larger than the time for a ready page, which is close to the average time per hit. Also, the average time to spent on a "nothing new" page is much less than average.

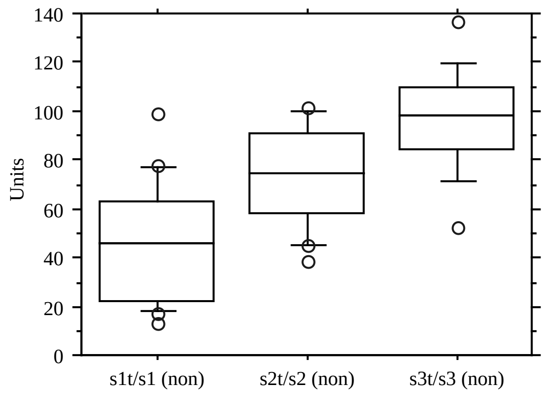


Figure 7: Average time per page for the case when the student navigate to a page using a non-annotated link .

What we observe on the [Fig. 7] isa "real value" of page state. Here students navigate to a selected page with a non-annotated link and without any warning about a page state. This data shows that the mechanism which determines different classes of pages works quite well. A page classified as "nothing new" can be read much faster (or just passed over) because it has no new information and a page classified as "not ready" is the most hard to understand because some background can be missed.

The data are quite different for the minority of pages (about 20% of hits) selected with annotatable link [Fig. 8]. What we observe on [Fig. 8] is a mixture of two effects: an effect of page state and an effect of annotation. This means that students of the ANS group noticed the annotations, may decide *apriori* to spend less time on "nothing new" pages and more on those annotated as "not ready". The effect of annotation clearly dominates in the case of "not-ready" pages. Those rare users who selected a page with full understanding that this page is not ready are willing to allocate significantly more time (ANOVA, p=0.012) for reading this page. Again, they understood how the system worked and trusted the integrity of the annotations.

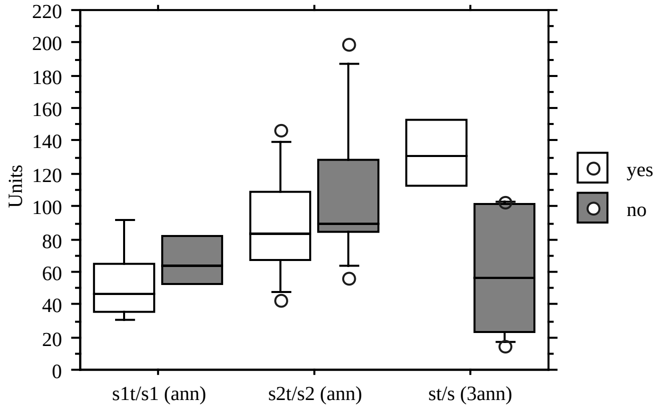


Figure 8: Average time per page for the case when the student navigate to a page using an annotatable link. The non-shaded "yes" bars are for the ANS group; the shaded "no" bars are for the non-ANS group

8 Discussion and Conclusion

These results suggest that ANS is a feature which is initially useful in improving comprehension for those new to a complex interface who are prepared to accept it. However, it adds another option to an interface: a cognitive overhead which may distract users from the content. This was reflected in the fact that the overall group who received ANS initially performed significantly worse in the knowledge tests. User model based link annotation seems to be of value to those that agree with it, those that accept and follow the annotations. Those learners who follow the annotated links are essentially in agreement with the cognitive model of the knowledge that the author, as expert, has placed on the content [Eklund 1995]. They take advantage of the fact that the content has been examined and structured for them, and they make use of both the implicit structure of the knowledge that the courseware embodies in its static form, as well as the individual link annotations which hint at the domain structure relative to the current path of the learner through the user model.

The experiment offers firm evidence that adaptive link annotation has an effect on student learning in an educational system. However, once users with minimal hits on the system were excluded, numbers in each of the two groups with and without annotations was a mere 8 and 9 respectively, and this is one of the severest limitations of the experiment. This study also suggests that the existence of ANS in the interface has an effect on the linearity of a user's path through the learning space, with those users experiencing link annotation being prepared to use those annotatable links more often than those who received no link annotation. As a result, their paths through the material were less linear, more exploratory, as they selected more "real links" and exhibited less use of the Continue button. In some ways this is hardly surprising, and again reflects the learner's trust of the system's annotations. If a link is annotated, a user has more confidence about the relevance of the material behind it than under a non-annotated link. In a non-annotated *InterBook* interface, the safest option for users was the repeated use of the Continue button.

A difficulty with the experimental procedure that was briefly described is that the majority of the navigation features were not annotated, so the difference in the interface for those students who did receive ANS was marginal. Another factor to be considered was the indexing of content and the subsequent authoring of the electronic textbook. This involves the allocation of a set of attributes for each domain node as both pre-requisite and outcome concepts, as described earlier. This indexing implies an ideal order in which to view content, represented by the continue button, as well as optional ways to view it, either recommended by green annotations or not recommended by those nodes annotated with red bullets. In the non-hypertext world, it is easy to flip the pages of a book in order, pausing only for those pages of interest. That way a reader may be sure that all the content has been examined, even if briefly. Similarly, it appears that one popular strategy was to follow the non-annotated continue link, and immediately move forward to the next node if the material on the current node is already known or of little interest. Students clearly felt that the continue link was the simplest and quickest way to review the material, thus following the domain structure imposed by both the textbook and the authoring process. This is a favourable outcome for adaptive curriculum sequencing, but not necessarily for adaptive link annotation.

More generally, it can also be argued that the lack of detail in the artificial world of the user-model somewhat trivialises the broad range of human responses and motives possible in learning. Since the start of the 1990s, this has been identified as an "intractable problem" [Self 1990], although more recent approaches give the learner access to this information as a goal-planning aid and learning-reflective device [Kay1997]. Is it possible to make a reasonable suggestion of the next best link for a user to follow with such a paucity of information about that individual? This revisits the problem of building practical intelligent tutors as discussed in the AI literature since the early 1980s. Even in well defined, highly organised and simplified domains, implementing a reliable system to account for user preferences, knowledge and idiosyncratic behaviour is highly problematic. In terms of the ramifications for informing the design of this experiment, to obtain a measurable favourable result for adaptive link annotation which confirms some of the interpretations of the study as presented above will require a simplification of the domain and a greater control of the variables earlier described. This may be difficult to achieve with an experiment situated in a real world teaching and learning context.

Acknowledgments

Many thanks to Elmar Schwarz of Tellux, Germany, for his advice about this research and for making available *InterBook* as an experimental vehicle, as well as Associate Professor Ken Sinclair and Dr. Mike Bailey of the Faculty of Education, University of Sydney, for their advice in the design and implementation of components of the empirical study discussed in this paper.

References

[Brusilovsky 1996] Brusilovsky, P.:"Methods and techniques of adaptive hypermedia"; User Modeling and User-Adapted Interaction, 6, 2-3 (1996) 87-129.

[Brusilovsky, Eklund & Schwarz 1998] Brusilovsky, P., Eklund, J., & Schwarz, E.: "Web-based education for all: A tool for developing adaptive courseware"; Computer Networks and ISDN Systems, 30, 1-7 (1998),291-300.

[Brusilovsky & Pesin 1994] Brusilovsky, P., & Pesin, L.: "An intelligent learning environment for CDS/ISIS users"; Proc. The interdisciplinary workshop on complex learning in computer environments (CLCE94), EIC, Joensuu, Finland (1994) 29-33, also apeared as electronic version, http://cs.joensuu.fi/~mtuki/www_clce.270296/Brusilov.html.

[Brusilovsky & Pesin 1998] Brusilovsky, P., & Pesin, L.: "Adaptive navigation support in educational hypermedia: An evaluation of the ISIS-Tutor"; Journal of Computing and Information Technology, (1998).

[Brusilovsky, Pesin & Zyryanov 1993] Brusilovsky, P., Pesin, L., & Zyryanov, M.: "Towards an adaptive hypermedia component for an intelligent learning environment"; In Bass, L.J., Gornostaev, J., & Unger, C. (Eds.), Human-Computer Interaction, Springer-Verlag, Berlin (1993) 348-358.

[Brusilovsky, Schwarz & Weber 1996] Brusilovsky, P., Schwarz, E., & Weber, G.: "ELM-ART: An intelligent tutoring system on World Wide Web"; In Frasson, C., Gauthier, G., & Lesgold, A. (Eds.), Intelligent Tutoring Systems, Springer Verlag, Berlin (1996) 261-269.

[de La Passardiere & Dufresne 1992] de La Passardiere, B., & Dufresne, A.: "Adaptive navigational tools for educational hypermedia"; In Tomek, I.(Ed.) Computer Assisted Learning, Springer-Verlag, Berlin (1992)555-567.

[Eklund 1995] Eklund, J.: "Cognitive models for structuring hypermedia and implications for learning from the world-wide web"; Proc. Ausweb95: The First Australian World-Wide Web Conference, Southern Cross University Press, Ballina, Australia (1995) 111-116, also apeared as electronic version,

http://elmo.scu.edu.au/sponsored/ausweb/ausweb95/papers/hypertext/eklund/index.html.

[Hohl, Bšcker & GunzenhŠuser 1996] Hohl, H., Bšcker, H.-D., & GunzenhŠuser, R.: "Hypadapter: An adaptive hypertext system for exploratory learning and programming"; User Modeling and User-Adapted Interaction, 6, 2-3 (1996) 131-156.

[Kay 1997] Kay, J.: "Learner know thyself: Student models to give learner control and responsibility"; Proc. ICCE97, International Conference on Computers in Education, Malasia, Kuching, Sarawak (1997) 18-26.

[Rubin 1996] Rubin, C.: "Macintosh Bible Guide to CarisWorks 4"; Peachpit Press, CA (1996)

[Self 1990] Self, J.A.: "Bypassing the intractable problem of student modelling"; In Frasson, C., & Gauthier, G. (Eds.), Intelligent Tutoring Systems: At the crossroads of artificial intelligence and education, Ablex Publishing, Norwood (1990) 107-123.

[Weber & Specht 1997] Weber, G., & Specht, M.: "User modeling and adaptive navigation support in WWW-based tutoring systems"; In Jameson, A., Paris,C., & Tasso, C. (Eds.), User Modeling, Springer-Verlag, Wien (1997) 289-300.

[Zhao, O'Shea & Fung 1993] Zhao, Z., O'Shea, T., & Fung, P.: "Visualization of semantic relations in hypertext systems"; Proc. ED-MEDIA '93, World conference on educational multimedia and hypermedia, AACE, Orlando, FL (1993) 556-564.