Journal of Universal Computer Science, vol. 14, no. 17 (2008), 2799-2818
submitted: 30/6/08, accepted: 27/9/08, appeared: 28/9/08 © J.UCS

A Spiral Model for Adding Automatic, Adaptive Authoring to Adaptive Hypermedia

Maurice Hendrix
(Department of Computer Science, The University of Warwick, Coventry, United Kingdom
maurice@dcs.warwick.ac.uk)

Alexandra Cristea
(Department of Computer Science, The University of Warwick, Coventry, United Kingdom
acristea@dcs.warwick.ac.uk)

Abstract: At present a large amount of research exists into the design and implementation of adaptive systems. However, not many target the complex task of authoring in such systems, or their evaluation. In order to tackle these problems, we have looked into the causes of the complexity. Manual annotation has proven to be a bottleneck for authoring of adaptive hypermedia. One such solution is the reuse of automatically generated metadata. In our previous work we have proposed the integration of the generic Adaptive Hypermedia authoring environment, MOT (My Online Teacher), and a semantic desktop environment, indexed by Beagle++. A prototype, Sesame2MOT Enricher v1, was built based upon this integration approach and evaluated. After the initial evaluations, a web-based prototype was built (web-based Sesame2MOT Enricher v2 application) and integrated in MOT v2, conforming with the findings of the first set of evaluations. This new prototype underwent another evaluation. This paper thus does a synthesis of the approach in general, the initial prototype, with its first evaluations, the improved prototype and the first results from the most recent evaluation round, following the next implementation cycle of the spiral model [Boehm, 88].

Keywords: Authoring; Adaptive Educational Hypermedia; CAF; Evaluation; Metadata; RDF; Semantic Desktop; Semi-automatic adding
Categories: M.5, M.6, M.7, M.8

1 Introduction

Authoring of adaptive hypermedia is a notoriously difficult endeavour [Celik et al., 06] although its results can be extremely valuable, generating, e.g., in the educational context, personalized (learning) experiences [Brusilovsky, 01]. A solution to this problem is to use as much automatically generated authoring as possible, instead of authoring by hand. There is some research into how to automate authoring in different ways [Cristea et al., 05], [Boticario et al. 07], [Martin et al. 07], [Vialardi et al. 07]. A good basis of this is to use already annotated resources, which can be automatically retrieved when necessary, as dictated by the authoring process. A rich source of information that we found can be exploited in this sense is the semantic desktop [Chirita et al., 05], [Semantic Desktop, 08]. In the semantic desktop, resources can be categorized by rich ontologies, and semantic links express various kinds of semantic relationships between these resources. For a document, for example, the semantic desktop stores not only a filename, but also information about *where* this document

was created, *when* and by *whom*, *which* of the colleagues sent it, and *how often* and *in what context* it was accessed. All these metadata are generated automatically, by the appropriate applications, and stored in an application independent way as RDF metadata [W3C RDF 04], in the user's personal data store. This rich set of metadata clearly makes it easier for the user or applications to (semi-)automatically retrieve appropriate material for different contexts: for example, when an author wants to select appropriate materials for a lecture. Of course, in the latter context, the author still has to create some basic lesson material, serving as a retrieval framework.

Previously, in [Hendrix et al., 06], we described the interaction and exchange of data between the Beagle++ environment [Beagle, 08], [Chirita et al., 06], which is an advanced search and indexing engine for the semantic desktop, generating and utilizing metadata information, and the adaptive hypermedia authoring environment MOT (My Online Teacher) [Cristea et al., 03], [MOT, 08], a sophisticated system for authoring personalized e-courses. This paper presents a synthesis of the approach, as well as initial evaluations of the of the Sesame2MOT Enricher v1 (converter) application, leading to the improved web-based Sesame2MOT Enricher v2 application and the most recent evaluation. The latter analyses the improved Enricher v2 that guides the interaction process between MOT and Sesame and performs the automatic authoring and population of the authoring environment with new resources (here, papers) from the semantic desktop.

The remainder of this paper is organized as follows. In section 2 we provide a brief scenario, outlining how automatic authoring could be used. Section 3 shortly introduces our approach to Adaptive Authoring of Adaptive Hypermedia, as well as the system setup, which implements a prototype of this approach. In section 4 we briefly describe our initial prototype. In Section 5 we describe the initial evaluations. In section 6 the results of these evaluations are discussed. After the evaluation, an improved web-based prototype was built. This prototype is discussed in section 7. In section 8 we discuss the evaluation setup for the evaluation of the improved prototype. In Section 9 we discuss the results of the last evaluation round. In section 10 we discuss related work. Finally, in section 11, we draw conclusions, discussing what these results mean for our approach in general and the prototype in particular.

2 Authoring and corresponding Learning Scenario

Consider Dr. Brown who prepares a new on-line course on early 19th century European history for 1st year history students at Warwick. The university allocates a limited amount of time for this. Adaptive Hypermedia is considered to result in a better learning experience [Brusilovsky, 01] and he wants to benefit from automatic help during authoring. Therefore he decides to use MOT [Cristea et al., 03], an *adaptive educational hypermedia authoring environment* with *adaptive authoring support.* This decision means that he would only have to author a standard basic course, as he usually does. Still, this costs him slightly more time, as he has to divide his course into conceptual entities and add semantic labels. After this, however, the authoring system can *automatically* enrich the course based on (pedagogical) strategies chosen by Dr. Brown. For example, the system can consider the version of

the course created by Dr. Brown as the version for *beginner*[1] students, which don't aspire at higher grades or deep knowledge on the subject. For *advanced* students, who want to achieve high grades, or deeper knowledge, the adaptive authoring system can use *semantic personal desktop search* to automatically find any existing scientific papers that are relevant to the current course on Dr. Brown's desktop. These scientific papers can be used as alternative or additional material to the main storyline of the static course. This mechanism builds upon a number of assumptions:

- as Dr. Brown is a specialist in the subject taught or at least has advanced knowledge of and a reasonable amount of interest in the subject. He will therefore at least read papers of interest on the subject, which are likely to be stored on his computer;
- research papers on the subject taught are considered as useful *extra* resources for *advanced learners*;
- Dr. Brown has been using Beagle++ Semantic Desktop System [Beagle, 08], [Chirita et al., 06] to store both papers and all relevant metadata automatically, in RDF format.

The storing of papers will have taken place over many years, and Dr. Brown may not know exactly where on his computer each article relevant to the current course is. However, the Beagle++ Semantic Desktop System can be used by the authoring tool to retrieve relevant papers found on Dr. Brown's desktop. He then selects the papers he wants[2], and fine-tunes the semantic labels for the chosen adaptation strategy.

Now, with relatively very little effort, Dr Brown has created an adaptive course, which will provide a richer learning experience for student.

3 The approach & system setup

In this section we introduce our method and system setup. As can be seen in Figure 1, Beagle++, the Semantic Desktop Environment used in our prototype, stores all metadata in the Sesame RDF database [Schlieder et al., 02]. All Beagle++ components that *generate* metadata (for example, the email, publication, web cache and file metadata generators) add the metadata to this database. All Beagle++ components which *use* metadata (for example, the search and indexing module, the ranking module or the browsing modules) retrieve their data from this repository, and, in some cases, write back new data (such as the PageRank value for documents or other resources). It is easy to accommodate additional modules in this environment by writing appropriate interface components, which read and write from this repository. This is what we have done for MOT [Cristea et al., 03], [MOT, 08]. We have focused

[1] **Please note** that the approach is not limited to a beginner- intermediate- advanced strategy, a teacher can decide to use any strategy he likes and set labels and weights accordingly. It is always a difficult trade-off between offering as rich a functionality as possible while keeping the interface as simple as possible. For this reason we have provided default settings for many options. The defaults will not always be ideal, but should provide a workable result. An other example that could be employed is, e.g., a resource *type*-based strategy, see: http://prolearn.dcs.warwick.ac.uk/strategies.html

[2] This extra selection step is necessary for privacy: Dr. Brown needs to approve which papers from his desktop are to be shown to the students.

on the semi-automatic addition of articles stored on the user's desktop to a MOT lesson [Cristea et al., 03]). This represents an instantiation of the concept of adaptive authoring: authoring that adapts to the author's needs. In MOT, the addition is done to an *existing* lesson. Based on pedagogic goals, the author can then process the data, by adding more information on the article after the conversion.

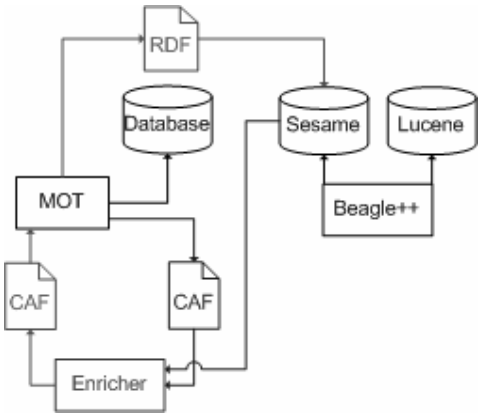

Figure 1: System overview

These additions can then be fed back into the RDF store, if necessary. We use CAF (Common Adaptation Format) [Cristea et al., 05], a system-independent XML exchange format, to simplify the transformation process from RDF to the MOT MySQL storage format.

3.1 Enrichment of the lesson and domain model

As MOT is mainly a tool for authoring educational (adaptive) material, the internal information structures are based on strict hierarchies (as advocated in [Cristea et al., 03]). When enriching the domain-models and lessons, we therefore aim at getting the right information in the right place in this hierarchy. To achieve this, the program first queries the Sesame database, using as search terms *title* and *keywords* of each domain concept found in the current existing lesson. The RDF query in the SeRQL [Broekstra et al., 04] language looks as follows:

```
select x from x {p} y where y like "*keyword" ignore case
```

Some alternative retrieval methods have been studied and implemented, as follows. For computing the mutual relevance between a given article and a given concept, in order to decide the appropriate place of articles in the concept hierarchy, we have developed two slightly different theoretical alternatives, called *Concept-Oriented* and *Article-oriented* methods. In the next sections we will discuss the ranking alternatives, as well as some options that influence the ranking results.

3.1.1 Concept-oriented relevance ranking method

The first relevance ranking method is computing relevance according to Equation 1:

$$rank(a,c) = \frac{|k(c) \cap k(a)|}{|k(a)|}$$

Equation 1 Concept oriented ranking

, where:

rank(a,c) is the rank of article a with respect to the current domain concept c;
k(c) is the set of keywords belonging to the current domain concept c;
k(a) is the set of keywords belonging to the current article a;
|S| = the cardinality of the set S, for a given set S.

This formula is *concept-oriented*, in the sense that articles 'battle' for the same concept: a given article is placed in the appropriate place in the hierarchy by it.

The actual implementation of the concept-oriented formula is not trivial for the RDF database that we have, as there is no keyword property available for retrieved objects (articles). Therefore, these keywords had to be extracted from the title property of the object. This means that the computation of the cardinality of the keywords, |k(a)|, can be interpreted in many ways. The easiest way is to set it to the number of words in the title property. Another option is to compute the number of relevant keywords (e.g., eliminating stop-words such as 'the', 'a', etc.) by connecting them to a given ontology or dictionary, such as WordNet. Both options were implemented; however, the current WordNet connection is very slow and thus inefficient.

3.1.2 Article-oriented relevance ranking method

A second implementation of the rank is given by Equation 2.

$$rank(a,c) = \frac{|k(c) \cap k(a)|}{|k(c)|}$$

Equation 2 Article oriented ranking

The equation shows how many of the keywords (shared by the article and the concept) are present in the concept. E.g., if a concept has less keywords than another one, but the keywords shared between article and concept are the same, the former concept will have a higher rank and 'win' the article. This formula shows thus to which of the concepts the article should be attributed. Therefore, this formula is article-oriented, in the sense that concepts 'battle' for the same article.

3.1.3 Sets versus multisets

Next, once the formula is chosen, there is another possible distinction to be made: we have chosen to implement two different versions for computing the cardinality of the intersection, one *set*-based (with intersection operation on sets, as defined above), and one with *multisets* or bags (and the respective intersection operation on bags).
In set theory, a *bag* can be formally defined as a pair *(A, m)* where A is some set and $m : A \rightarrow N$ is a function from A to the set N = {1,2,3,...} of (positive) natural

numbers. The set A is called the underlying set of elements. For each a∈A the multiplicity (i.e., number of occurrences) of a is the number *m(a)*.

The intersection can then be defined as: $(A \cap B, f(A \cap B))$, where: $f(x)=\min\{m(x \in A), m(x \in B)\}$. For example: *{(a,2),(b,1)}* ∩ *{(a,3),(x,1)}* = *{(a,2)}*

The reason to use bags instead of sets is that the number of times keywords appear in certain texts can be relevant in itself (not just *which* keywords). A text containing a greater number of occurrences of a specific keyword could be a better match for that keyword than a text with only one occurrence of the respective keyword.

3.1.4 Allowing duplicates or not

The same resource may be relevant in more than one place within the hierarchy. In that case, the resource will be added to the place where it has the highest relevance, by default. If there are more places in the hierarchy with a value equal to the highest relevance, the one with the higher position in the tree wins. If, however, there are siblings with the same position in the tree, and with the (same) highest relevance, a decision has to be made: either to allow duplicates, or to select randomly one of the candidate sibling concepts and allocate the resource to it. The '*allow duplicates*' option in the Sesame2MOT Enricher v1 application controls this particular option.

3.1.5 Adding meta-data as separate concepts or as attributes

The retrieved metadata also has a structure. For example, a retrieved paper might have a *location* it was presented at and a *year* it was presented in. This metadata can be added either as attributes of the new *article* concept in MOT, or as a set of new sub-concepts, with their own attributes. As we can see in Figure 2, with the '*Add meta-data as separate concepts*' option, the author can select between these two possibilities in the Sesame2MOT Enricher v1 application.

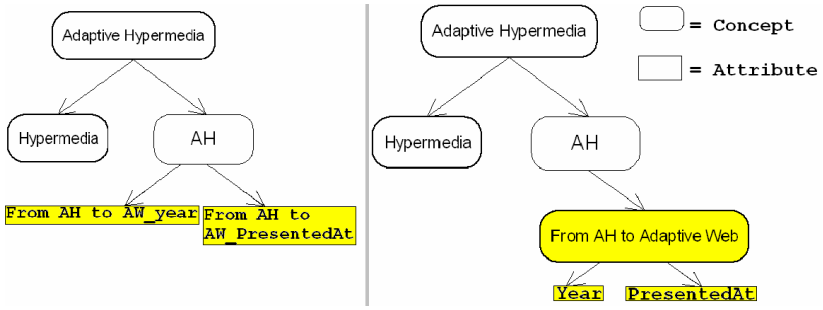

Figure 2: Right side, result with 'add metadata as separate concepts', left: without

4 Initial prototype

The initial prototype was built as a stand-alone Java application, due to the fact that it was supposed to be placed on the same machine as the semantic desktop application,

and it needed the access rights to the articles. The processing steps to be done by authors were as follows: step 1) download a CAF XML file from the MOT authoring system, then step 2) load this CAF XML file into the prototype, step 3) add resources as described in the previous section, step 4) save the CAF XML file and step 5) upload this new, updated CAF XML file to the MOT authoring again.
The interface interaction with users is depicted in Figure 3. The first two steps, to be performed in the MOT authoring system, are not shown, as the focus is on the Sesame2MOT Enricher v1 application. Figure 3 displays the resource addition in step 3. Authors can choose between concept-oriented versus article-oriented processing, between allowing or not duplicates among siblings, between computing the resource keywords as sets or as multi-sets, and between adding meta-data about articles as separate concepts, or as attributes of the matching domain concept. All these options conform to the theoretical description in the previous section. Moreover, the author has the option to change the default way that the new resources will be labelled: they would by default be labelled as advanced, with a weight of 50. If the author would like to later use a different educational adaptation strategy with this course, he could change the labelling and weighting manually in the MOT authoring environment.

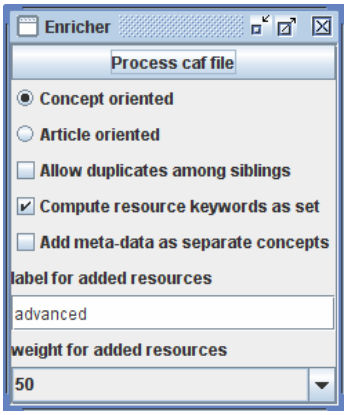

Figure 3: CAF loading; options and ranking method selection screen

After specifying the desired options for the matching mechanism between articles and domain concepts, the list of possible resources to add is displayed to the author, as depicted by Figure 4. This represents the second part of step 3 in the processing sequence. This step allows the author to overview the automatic matching process and hands him the ultimate decision about which papers to actually include in the final course. This allows for a checking stage, where unsuitable documents are filtered out (for instance, the exam paper with solutions on the topic of databases might appear as very suitable for the database course, but it still is not appropriate to show to the students). For such reasons, a fully automatic processing at this stage is inappropriate. By using semi-automatic processing, the teacher keeps the overall control.

After the author has decided about which articles to include, he would save them by pressing the ‘save caf & zip file’ button (Figure 4), and then deciding where on his

computer to save the exported CAF XML file. These actions represent step 4 in the processing.

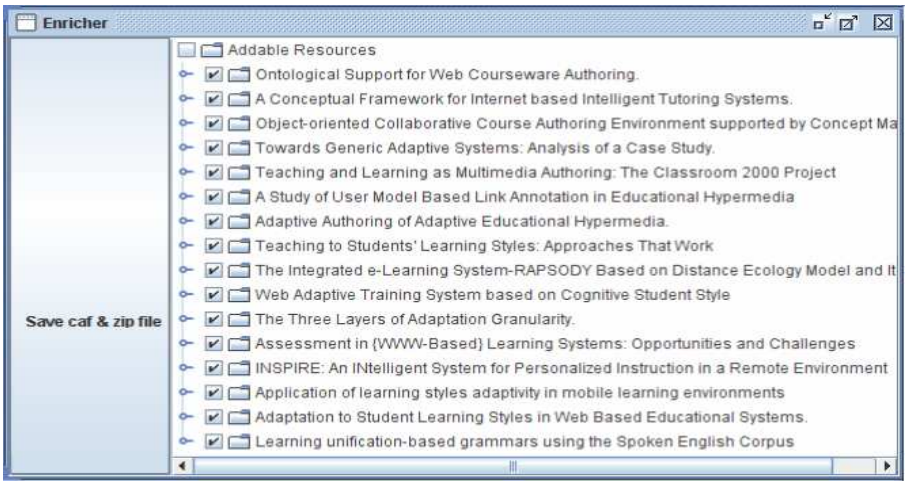

Figure 4: Resource selection screen

5 The evaluation setup

5.1 Preliminary evaluation

The initial evaluation of the conversion process and the Sesame2MOT Enricher v1 application performing took place in two steps: the first step was a small-scale qualitative experiment with about 4 PhD candidates in Educational Technology attending the IMPDET course organized by the Joensuu University, Finland. This evaluation was based on the *think-aloud* method (see more information in [Hendrix et al., 06]). The second step was of a larger scale, and contains therefore a larger amount of quantitative evaluation results, although qualitative information was also sought.

5.2 Second evaluation setup

This second evaluation step was conducted at the Politehnica University of Bucharest in January of 2007. It took place within an intensive two-week course on 'Adaptive Hypermedia and The Semantic Web'[3], which was delivered as an alternative track to the regular Intelligent Systems course. The students were 4th year undergraduates in Engineering studies and 2nd year Master students in Computer Science, from the English-language stream. Firstly, basic knowledge on Adaptive Hypermedia and Semantic Web was addressed. The first course week was dedicated to theory, and finished with a theoretical exam. Out of the initial 61 students, only the students with

[3] Please note that the lecture about the prototype and the theory behind it took about 1.5 hours. This training part was similar for all courses and interaction sessions used for the evaluation.

a satisfying knowledge of the theory where selected to continue with the practical part of the alternative track. The rest returned to the main stream course. The 33 students that passed the theory worked with the actual systems described in section 2. This selection process ensured that the evaluations are performed by people who understand at least what they are aiming at. After these experiments, they where requested to submit a number of questionnaires, to answer both generic and specific issues regarding the automatic generation of adaptivity and personalization via the methods described in section 2.

5.3 Hypotheses

We based our evaluation firstly on a number of generic, high level hypotheses:

- **v1 H 1.** The respondents enjoyed working as authors in the system.
- **v1 H 2.** The respondents understood the system.
- **v1 H 3.** The respondents considered that theory and practice match.
- **v1 H 4.** The respondents considered the general idea of Adaptive Authoring useful.

We refined these into more specific, lower granularity hypotheses (see Table 1), which ultimately generated our questions for the questionnaires. To explain the construction of the sub-hypotheses, let's take, for instance, hypothesis 3. There, we check the matching between theory and practice, i.e., between theory and the implementation. For the Sesame2MOT Enricher v1 application, from a theoretical point of view, we have defined different ranking methods and other options, such as allowing duplicates or not between the imported articles, etc. These have been implemented as options for the user to select, and therefore, in this particular case, matching theory and practice means that these methods render different results, firstly; and secondly, that these different results should be just as the theory has predicted. Therefore, sub-hypothesis 3.4, and its own sub-hypotheses, 3.4. 1, 3.4.2 and 3.4.3 emerged. As said, the hypotheses and sub-hypotheses feed into and determined the questions. The questions, unlike the hypotheses, were neutralized to the extent it was possible. Also, additionally to our specific questions, the SUS (system usability scale) questionnaire was used, which is a standard questionnaire especially aimed at removing bias, by alternating negative and positive questions. For brevity reasons, the SUS questionnaire results are not shown here (see, e.g.,[Hendrix et al. 07]).

6 Initial Evaluation Results and Discussion

In this section we present and discuss only the results from one of our initial evaluation, which focussed on the Sesame2MOT Enricher v1 application. We will first shortly discuss the quantitative feedback in general terms and then outline the numerical feedback obtained. We will show whether the hypotheses mention in section 5.2 have been confirmed.

The qualitative feedback showed that the stand-alone nature of the Sesame2MOT Enricher v1 application was quite confusing and an integrated version would be needed. Users considered the separate access to the different systems cumbersome, and many installation steps and manual data interchange steps were considered time

consuming and frustrating. It was also pointed out that, whilst working fine in isolation, when a whole group used the prototype at the same time, the connection between the Java application and the Sesame server was unreliable. We note that the initial idea was that people would use the Sesame server containing their own semantic-deskop, on their own desktop. The issue was improved considerably with the integrated version, as the two servers (running the authoring system and the Enricher application, respectively) were no longer far apart. A final problematic issue was the selection of appropriate weights and labels. Selecting these was found to be not very straightforward, and the application did not allow for individual changes of weights and labels of articles. These problems were addressed in the next, improved implementation, where users can choose to use the recommended labels & weights with each strategy or not, and then later on still have the possibility to change individual or groups of labels and weights. Moreover, the new prototype also features a strategy selection, which influences the weights and labels selected.

For testing our hypotheses we have used, where possible, numerical averages, and tested their significances with the help of Student's T-test. We assumed a confidence of 95%. Thus, a hypothesis can be confirmed, if the mean is $M>0$ and the probability is $P <0.05$. In order to obtain numerical averages, we mapped the multiple-choice answers follows: '*Yes*' was mapped to 1, '*no*' to -1 and '*mostly*' to 0. Hence the average was always 0 and the T-test was applied by comparing against the neutral result of 0. Below we present a table with each hypothesis, T-test results (T value, degrees of freedom Df, Mean M, probability P) and whether the results show that it was confirmed or not. The main hypotheses are shown in bold. Their result is obtained by combining the results of the sub-hypotheses.

As we have seen, most hypotheses have been confirmed based on the current data. The Sesame2MOT conversion is indeed considered useful and in line with the theory. Its options are understood. Respondents agreed strongly with most of our hypotheses, with all means above zero. Looking at the ones with lower scores, such as concept-oriented and article-oriented method, as well as computation of resource as set, they were less sure in their statements. This is probably due to the fact that they did not work with these options enough. This shows that more targeted evaluations may be necessary to establish without a doubt the acceptance rate of these features.

Nr.	Hypotheses	T	Df	M	P	Con-firmed (M>0; P<0.05)
v1 H1	**The respondents enjoyed working as authors in the system**	2.709	31	0.438	0.011	Yes
v1 H2	**The respondents understood the system.**					**Yes**
v1 H 2	Respondents *understood how all works*; in particular:					Yes
v1 H 2.1	Concept oriented formula	4.458	31	0.625	0.000	Yes
v1 H 2.2	Article oriented formula	3.788	31	0.563	0.001	Yes
v1 H 2.3	Allowing duplicates/ not	10.063	31	0.875	0.000	Yes
v1 H 2.4	Computing resources as set/ not	5.271	31	0.688	0.000	Yes
v1 H 2.5	Adding meta-data as separate concepts	6.313	31	0.750	0.000	Yes
v1 H 3	**The respondents considered that theory and practice match.**					**Yes**
v1 H 3.1	The two ranking methods (concept-, article-oriented) do deliver different results.					Yes
v1 H 3.1.1	Concept Oriented delivers a different result from Article Oriented [4]	6.313	31	0.750	0.508	Yes
v1 H 3.1.2	Article Oriented delivers a different result from Concept Oriented	6.313	31	0.750	0.508	Yes
v1 H 3.2	Ranking methods (concept-, article-oriented) *are in line with the theory*. In particular:		31			Yes
v1 H 3.2.1	Concept Oriented	2.252	31	0.375	0.032	Yes
v1 H 3.2.2	Article Oriented	2.709	31	0.438	0.011	Yes
v1 H 3.3	The different *options influence the result*; in particular:		31			Yes
v1 H 3.3.1	Allowing duplicates	7.760	31	0.813	0.000	Yes
v1 H 3.3.2	Computing resources as set/ not	3.215	31	0.500	0.032	Yes
v1 H 3.3.3	Adding meta-data as separate concepts/ not	6.313	31	0.750	0.000	Yes
v1 H 3.4	The results of the conversion *are in line with the theory*; in particular:					Yes
v1 H 3.4.1	The two ranking methods	see 3.3.1.2	see 3.3.1.2	See 3.3.1.2	see 3.3.1.2	(see 5.3.1.2)
v1 H 3.4.2	Allowing duplicates/ not	7.760	31	0.813	0.000	Yes
v1 H 3.4.3	Computing resources as set/ not	2.252	31	0.375	0.032	Yes
v1 H 3.4.4	Adding meta-data as separate concepts/ not	4.458	31	0.625	0.000	Yes
v1 H 4	**General idea useful**	**15.000**	**31**	**0.938**	**0.000**	**Yes**

Table 1: Sesame2MOT Conversion hypotheses results

[4] Note that logic dictates that v1 H3.1.1 and v1 H3.1.2 should have the same result. They were put in different questions formulated slightly differently to check for bias.

7 Improved web-based prototype

The new, improved prototype is web-based. It is integrated with the 'My Online Teacher' system. Therefore, the author does not need to deal with CAF files anymore. CAF files are still used behind the scenes. The author accesses the new prototype by clicking the ***enrich*** link in the ***Goal Model view*** of MOT as shown below.

Figure 5: New prototype is accessed via enrich link in Goal Model view

This link will take the author to a screen, shown in Figure 6, where he can select the ranking method and the other options as mentioned in section 3, similarly to the old prototype. The screen does feature however a new option of strategy selection. Users can choose to select a strategy and let the weights & labels be recommended based on their choice of strategy. The CAF file does not need to be loaded. As the web-based system knows which Goal Map the author accessed previously and has access to the CAF convertor, it can generate the CAF file without intervention from the user.

The next step features an improved version of the resource selection screen (Figure 7). The checkboxes allow deciding which of the recommended resources are to be added. On the right hand side there is a small outline of the Goal Map hierarchy. If the author clicks on one of the titles of resources it will highlight (in yellow) the place in the hierarchy where the resource will be added to when the checkbox is checked. In this way, more visual information and a sense and feel of the structure and placement of new resources is given to the authors, as compared to the initial prototype. Moreover, the form at the bottom allows the author to change weights and labels for individual resources (for all resources currently highlighted). The resource selection screen is shown in Figure 7.

Finally, after the author has selected the resources, the process as described in Figure 1 is completed. As the web-based Sesame2MOT Enricher v2 application has direct access to the import functionality for CAF files in MOT, there is no need for the author to manually upload this CAF file. The application does however retain the option for the author to download the modified CAF file. The application does however retain the option for the author to download the modified CAF file. This can be useful, for example, for use elsewhere or backup purposes.

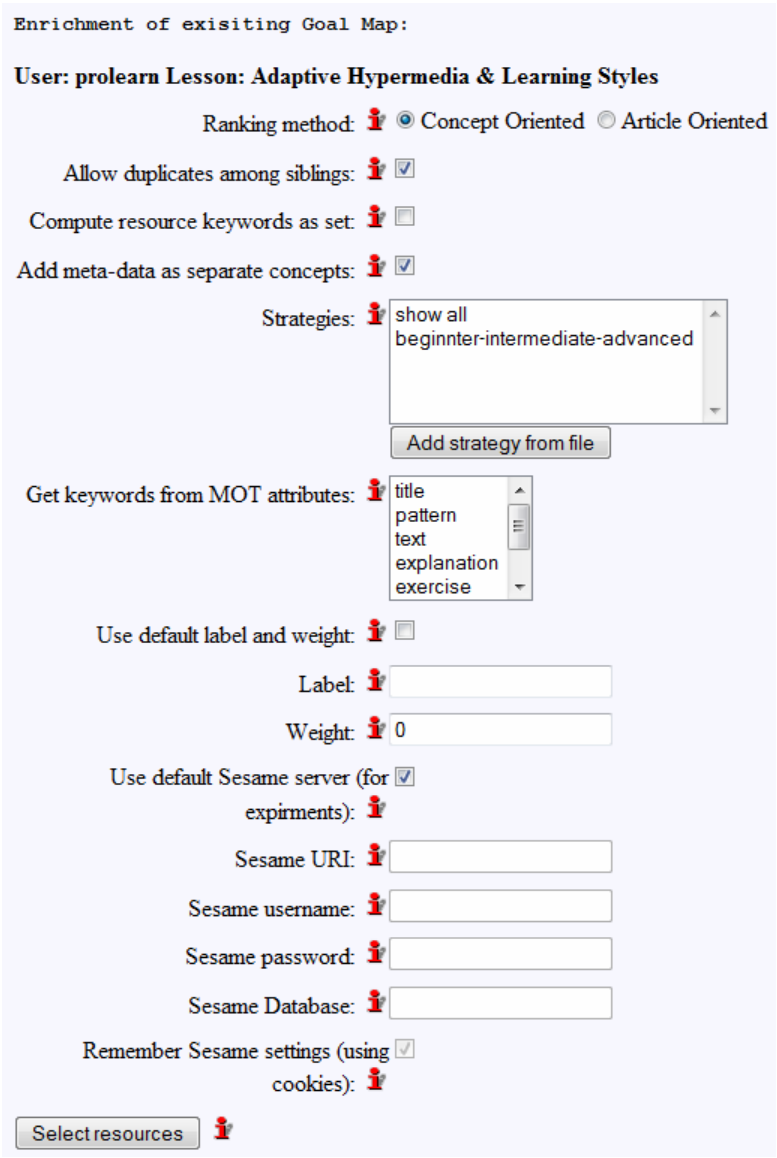

Figure 6: Options and ranking method selecting screen

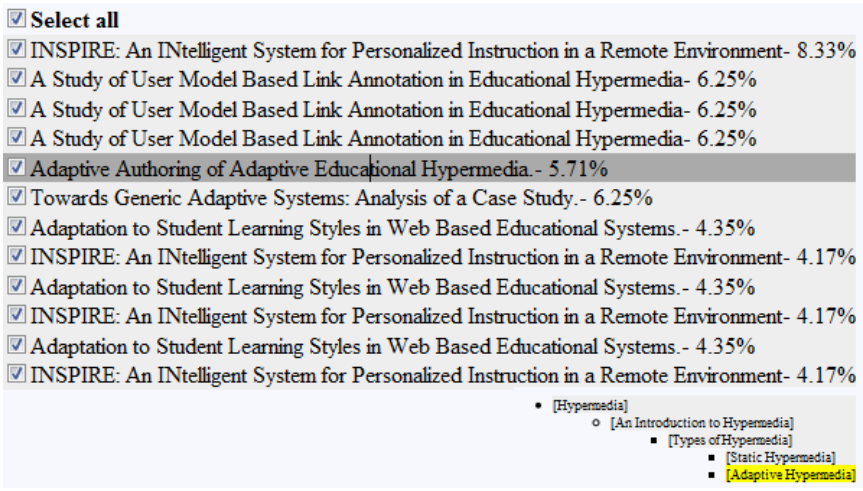

Figure 7: Resource selection

The deployment screen allows the author to deploy the new enriched Goal Map. The option of deploying to the AHA! delivery engine is similar to importing the modified CAF file in the previous prototype [Cristea et al., 05], but requires less effort from the author. There is also an additional option to *deploy to AHA!*. This offers a way to directly convert the created course into a lesson in an AHA! System. For this, the application requires some additional information on where to find AHA!. The deployment screen implementing the described step is shown below.

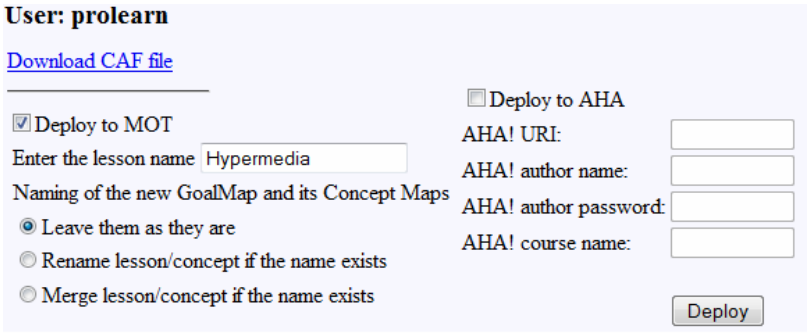

Figure 8: Deployment screen

8 Final Evaluation setup

8.1 General Evaluation Setup

The third and final evaluation was again conducted with another batch of students from the 'Politehnica' University of Bucharest, Romania and focused on the new

web-based Sesame2MOT Enricher v2 application. It was conducted in May 2008 with a group of thirty students studying in a "Web Programming" course[5], partially delivered via distance learning. Students collaborated in the creation of new content in MOT and answered our questionnaire. The students were enrolled in the 3rd year of Computer Science. Note that the web-based prototype was particularly suitable because the course was about web programming and was partially delivered online. The course covered, among other things, the basics of Adaptive Hypermedia and the theory behind out system setup. Directly after the lecture on our system setup students were requested to fill a questionnaire.

8.2 Hypotheses

We again based our final evaluation firstly on a number of generic, high level hypotheses, follows:

- **v2 H 1.** The students enjoyed working as authors in the system.
- **v2 H 2.** The students (think they) understand the new web-based Sesame2MOT Enricher v2 function.
- **v2 H 3.** Students think the interface of the web-based Sesame2MOT Enricher v2 application is clear.
- **v2 H 4.** Students perceive the web-based Sesame2MOT Enricher v2 application as being fast enough[6].
- **v2 H 5.** Students notice that the two ranking methods render deliver different results.
- **v2 H 6.** Students notice that the results of the respective ranking methods are in line with the theory (so did what the students expected).
- **v2 H 7.** Students understand the use and theory behind the ranking methods.
- **v2 H 8.** Students understand the option 'Compute resource keywords as set'.
- **v2 H 9.** Students notice that the option 'Compute resource keywords as set' influences the results.

9 Final Evaluation results and discussion

Again, due to lack of space, we present and discuss here only the numerical results from one of our questionnaire that focussed on the web-based Sesame2MOT Enricher v2 application. For testing our hypotheses, we have again used, where possible, numerical averages, and tested their significances with the help of Student's T-test. This time we had 3 possible answers for each question *'yes'*, *'no'* and *'no answer'*. We assumed *'no answer'* to be neutral and mapped the answers to 1, -1 and 0 respectively. Note that this means the average is 0 and tests were therefore done against the average (neutral) value of 0. Assuming a confidence of 95%, a hypothesis can be confirmed, if the mean is $M>0$ and the probability is $P<0.05$. A table with

[5] Please note the lecture on the prototype and the theory behind it took about 1.5 hours. This part was similar for both courses used for the evaluation.

[6] Please note that 'fast enough' reflects here the *perceived performance* for the user, to be distinguished from the actual, measurable performance.

each hypothesis, T-test results (T value, degrees of freedom Df, Mean M, probability P) and whether the results show that it was confirmed or not is displayed below.

Nr.	**Hypotheses**	**T**	**Df**	**M**	**P**	**Confirmed (M>0; P<0.05)**
V2 H 1	The students enjoyed working as authors in the system.	4.219	22	0.52174	0.000	Confirmed
V2 H 2	The students (think they) understand the new web-based Sesame2MOT Enricher v2 application.	3.219	22	0.39130	0.004	Confirmed
V2 H 3	Students feel the interface of the web-based Sesame2MOT Enricher v2 application is clear.	2.152	22	0.34783	0.043	*Not Confirmed*
V2 H 4	Students think the web-based Sesame2MOT Enricher v2 application is fast enough.	0.569	22	0.08696	0.575	*Not Confirmed*
V2 H 5	The students notice that the two ranking methods do deliver different results.	5.348	22	0.56522	0.000	Confirmed
V2 H 6	The students notice that the results of the respective ranking methods are in line with the theory (so did what the students expected).	4.899	22	0.52174	0.000	Confirmed
V2 H 7	The students understand the use and theory behind the ranking methods.	2.598	22	0.39130	0.016	*Not Confirmed*
V2 H 8	The students understand the option 'Compute resource keywords as set'.	3.425	22	0.52174	0.002	Confirmed
V2 H 9	The students notice that the option 'Compute resource keywords as set' influences the results.	2.237	22	0.39130	0.036	*Not Confirmed*

Table 2: Final evaluation results, web-based Sesame2MOT Enricher v2 application

We have seen that students enjoyed working with the new web-based Sesame2MOT Enricher v2.0 application. And after our explanations and working with it, they understand it. Whether the interface is clear and whether the applications works fast enough could not be confirmed. The performance bottleneck was found to be the connection to the Semantic desktop setup.

The students noticed the difference between the ranking methods and found them to be in line with the theory; although we could not establish whether they understood the actual use of the different ranking methods (although P is low so could be

acceptable for different confidence requirements. On the other hand, students understood the option 'Compute resource keywords as set' but we could not establish if they noted the difference. This may also be due to the fact that the examples they were told to use did not have many instances of multiple keywords occurrence in MOT.

We have furthermore experienced, whilst demonstrating the application to other fellow PhD students, that without the basic theoretical knowledge, the interface seems a little complicated. Balancing a high level of functionality and a high level of usability are often contradictory goals, and clearly, compromises have to be accepted. In the second prototype this was partially addressed by offering standard settings for most options. Also, a list of strategies was introduced, where the weights & labels would depends on the chosen strategy (unless the author chooses to change them). However, for the future, we recommend putting these options in a *settings* screen so that they don't confuse beginning authors.

Please also note that the populations used in the experiments were very different, both in terms of age and experience (PhD versus Master students and 3rd year students) as well as in terms of application areas (Educational Technology versus Engineering and Computer Science). In this way, we could obtain information about the tool usage from various potential users. Further work especially is concentrating on combining Semantic Web techniques with Social Web and Web 2.0 techniques, allowing students to be both learners and authors (contributors) to their learning environment. For this purpose, it is useful to estimate how all these types of users can deal with the system.

10 Related Work

Authoring of adaptive hypermedia is notoriously difficult work [Brusilovsky, 01]. Research on improving the process of adaptive hypermedia authoring ranges from ontology-based authoring [Martin et al. 07], to integrating standards and their representations [Gutierrez 07], [Boticario et al. 07], using data mining techniques [Vialardi et al. 07], web services [Meccawy et al. 06], interfacing techniques between authoring systems [Cristea et al., 04], and adaptation languages [Cristea et al., 05]. Moreover, for all the directions above, there are developments towards graphical authoring tools [Hendrix et al. 08], [Freire et al. 05], to aid the complex task of the authoring of (pedagogical) adaptation strategies.

Our work also relates to efforts by [Chirita et al., 06] regarding searching and ranking within RDF metadata. The difference is that we are comparing resources, from the Semantic Desktop setup, in RDF metadata with resources in the authoring environment which are not currently stored in a semantic way. [Dolog et al., 03] proposes a complete semantic-based approach, in which the user model would be stored in RDF and standards like [LOM] (Learning Object Metadata) would be used for describing the learning resources.

Our work has provided a new approach to exploit a Semantic Desktop [Beagle, 08], [Semantic Desktop, 08] for aiding Adaptive Hypermedia authors. Though the approach provides content, it also helps in the authoring of adaptation, as labels and weights can be provided for the retrieved resources to offer them as content-alternatives, according to a strategy chosen by the author.

11 Conclusions

In this paper we have briefly reviewed an authoring environment for personalized courses, as well as a Sesame2MOT Enricher mechanism and prototype based on Semantic Desktop technology. The paper describes the theoretical considerations for the implementation that uses data and relations stored by a semantic desktop environment in order to (semi)-automatically enrich adaptive course and then, in parallel, the evaluation of these considerations as well as of the prototype. From the three evaluation steps performed so far the promising result is that the theoretical concept of *Adaptive Authoring of Adaptive Hypermedia* was perceived as useful. However, the actual implementation has always room for improvements. We have gained some important feedback into possible improvements to the Sesame2MOT Enricher application itself, as described in section 6. Respondents in our experiments pointed out that the integration is currently not optimal and the user interface needs enhancements. Based upon some initial evaluations we developed a web-based prototype and ran a third series of evaluations. These evaluations showed that students enjoyed working with the new web-based Sesame2MOT Enricher v2.0 application, they mostly understood it and, importantly, they noticed the difference between ranking methods and confirmed that the ranking is in line with the theory. We did not get a conclusive answer on the interface and performance. We feel that the interface may need careful explanations and the performance could be improved, especially in terms of the connection to Sesame.

Concluding, we claim that this research has shown new ways in which the complex problem of authoring for personalization and adaptivity can be approached, by exploiting existent resources and introducing (semi)-automatic steps in the authoring process. In such a way, adaptation is performed at the early stage of authoring, and is not just an end product of authoring. This is definitely not the end of the way. Other means can be exploited to help authors in their tasks, and adapt to their needs. In previous research [Cristea et al., 03], we have shown that authoring adaptation can be performed by automatically adding relatedness relations between concepts in a domain map. This current research is bolder, in that it allows for adding both content and relations. We hope we have thus paved the way for further researches in this exciting area of Adaptive Authoring of Adaptive Hypermedia.

Acknowledgements

We would like to thank the Socrates mobilities exchange program at The University of Warwick, Politehnica University of Bucharest and Prof. Dr. Paul Cristea for giving us the opportunity to conduct our user study. This work has been performed in the framework of the IST project IST-2007-215434 GRAPPLE which is partly funded by the European Union. This work is based on findings from the ALS project 229714-CP-1-2006-1-NL-MPP and the PROLEARN Network of Excellence.

References

[Beagle, 08] Beagle++ website, viewed 13 May 2008, http://beagle2.kbs.uni-hannover.de/. 2008

[Boehm, 88] B.W. Boehm, A Spiral Model of Software Development and Enhancement, IEEE Computer, v.21 n.5, p.61-72, May 1988. 1988

[Boticario et al. 07] Boticario, J.G., Santos, O.C., A dynamic assistance approach to support the development and modelling of adaptive learning scenarios based on educational standards, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, Workshop, User Modelling, UM 2007, User Modelling, Corfu, Greece. 2007

[Broekstra et al., 04] Broekstra, J., Kampman, A., SeRQL: An RDF Query and Transformation, the International Semantic Web Conference, ISWC 2004, Florida USA. 2004

[Brusilovsky, 01] Brusilovsky, P., Adaptive hypermedia, User Modelling and User Adapted Interaction, Ten Year Anniversary Issue, 11 (1/2), 87-110. 2001

[Celik et al., 06] Celik, I., Stewart, C., Ashman, H. Interoperability as an Aid to Authoring: Accessing User Models in Multiple AEH Systems, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, Adaptive Hypermedia and Adaptive Web-based Systems, AH 2006, Dublin Ireland. 2006

[Chirita et al., 06] Chirita, P.-A., Costache, S., Nejdl, W., and Paiu, R. Beagle++: Semantically Enhanced Searching and Ranking on the Desktop. Proceedings of the 3rd European Semantic Web Conference, Budva, Montenegro. 2006

[Chirita et al., 05] Chirita, P.-A., Gavriloaie, R., Ghita, S., Nejdl, W., and Paiu, R., Activity-Based Metadata for Semantic Desktop Search. Proceedings of the 2nd European Semantic Web Conference, Heraklion, Crete. 2005

[Cristea et al., 03] Cristea, A. I., De Mooij, A.,Adaptive Course Authoring: My Online Teacher. IEEE International Conference on Telecommunications ICT'03, Papeete, French Polynesia. 2003

[Cristea et al., 05] Cristea, A.I., Smits, D., and De Bra, P., Writing MOT, Reading AHA! - converting between an authoring and a delivery system for adaptive educational hypermedia -, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, International Conference on Artificial Intelligence in Education AIED'05, Amsterdam, The Netherlands. 2005

[Cristea et al., 05] Cristea, A., Stewart, C., Automatic Authoring of Adaptive Educational Hypermedia, book chapter in "Web-Based Intelligent e-Learning Systems: Technologies and Applications", Book chapter in Zongmin Ma. IDEA Publishing group. 2005

[Cristea et al., 05] Cristea, A.I., Stewart, C., Ashman, H. and Cristea, P., Evaluation of Adaptive Hypermedia Systems' Conversion, Hypertext, HT'05, Salzburg, Austria. 2005

[Cristea et al., 04] Cristea, A., Verschoor, M., The LAG Grammar for Authoring the Adaptive Web, International Conference on Information Technology, ITCC'04, Las Vegas, US, IEEE. 2004

[Dolog et al., 03] Dolog, P., Gavriloaie , R., Nejdl , W., Brase , J., Integrating adaptive hypermedia techniques, World Wide Web Conference, WWW03, Budapest, Hungary. 2003

[Freire et al. 05] Freire, M. & Rodríguez, P., Comparing Graphs and Trees for Adaptive Hypermedia Authoring , Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH , International Conference on Artificial Intelligence in Education, AIED'05, Amsterdam, The Netherlands,. 2005

[Gutierrez 07] Gutierrez, S., Authoring of Adaptive Sequencing for IMS-LD, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, User Modelling, UM 2007, Corfu, Greece. 2007

[Hendrix et al. 08] Hendrix, M., De Bra, P., Pechenizkiy, M., Smits, D., Cristea, A., Defining adaptation in a generic multi layer model:CAM: The GRAPPLE Conceptual Adaptation Model, to appear, European Conference on Technology Enhanced Learning, EC-TEL 08, Maastricht, The Netherlands. 2008

[Hendrix et al. 07] Hendrix, M., Cristea, A., A qualitative and quantitative evaluation of Adaptive authoring of Adaptive Hypermedia, EC-TEL 2007, Second European Conference on Technology Enhanced Learning, Crete, Greece. 2007

[Hendrix et al., 06] Hendrix, M., Cristea, A.I., Nejdl, W., Authoring Adaptive Learning Material on the Semantic Desktop, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, Adaptive Hypermedia and Adaptive Web-based Systems, AH 2006, Adaptive Hypermedia Dublin Ireland. 2006

[LOM] LOM, Learning Object Metadata, IEEE-LTSC, http://ltsc.ieee.org/wg12/

[Meccawy et al. 06] Meccawy, M. Stewart, C. and Ashman, H., Adaptive Educational Hypermedia Content Creation: A Web Service based Architecture, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, Adaptive Hypermedia and Adaptive Web-Based Systems, AH2 006, Dublin, Ireland. 2006

[Martin et al. 07] Martin, B., Mitrovic, A., Suraweera, P., Domain Modelling with Ontology: A Case Study, Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia, A3EH, User Modelling, UM2007, Corfu, Greece. 2007

[MOT, 08] MOT homepage, viewed 13 May 2008, http://prolearn.dcs.warwick.ac.uk/mot.html

[W3C RDF 04] Resource Description Format, W3C Recommendation, http://www.w3.org/RDF/. 2004

[Schlieder et al., 02] Schlieder, T., Naumann, F., Approximate tree embedding for querying xml data, Heterogeneous and Distributed Information Retrieval, ACM SIGIR Workshop on CML and Information Retrieval. 2002

[Semantic Desktop, 08] Semantic Desktop, viewed 13 May 2008, http://www.semanticdesktop.org/. 2008

[Vialardi et al. 07] Vialardi, C., Bravo, J., Ortigosa, A., Empowering AEH Authors Using Data Mining Techniques, A3EH, Workshop, User Modelling, UM 2007, User Modelling, Corfu, Greece. 2007