Ontology-based Interface Adaptivity in Web-based Learning Systems

Arturo Peñarrubia, Antonio Fernández-Caballero and Pascual González
Universidad de Castilla-La Mancha, Instituto de Investigación en Informática de Albacete
Laboratory of User Interaction and Software Engineering (LoUISE), Albacete, Spain
{arpebla, caballer, pgonzalez}@info-ab.uclm.es
Federico Botella, Antonio Grau and Oscar Martínez
Universidad Miguel Hernández de Elche, Operations Research Center, Elche, Spain
{federico, a.grau, oscar.martinez}@umh.es

Abstract

Along this document, we expose how user interface adaptivity principles may contribute to improve present web learning systems. If we are able to establish some student features, based on the documents he has consulted, we can easily guide him through our web site. In order to collect that information, the web site considers the metadata associated with each document. Once this information has been obtained, the interface will recommend the site documents that are the most tailored to the student's interests. Our proposal consists in mixing the main ideas inherent to user profiling, semantic web, metadata, and ontologies. The framework supporting the initiative is also introduced in this paper.

1. Introduction

Recently, Human Computer Interaction (HCI) is awakening a growing interest from different perspectives, and especially from Software Engineering [3], [9]. The need to facilitate access to more and more complex systems is causing an unprecedented development in this research area. And especially web-based applications, as they must manage a big quantity of data with complex structures due to the wideness and heterogeneity of the Internet, are those which require a most important treatment with regard to interaction.

It is not a trivial task for a user who accesses such a large and complex information system to easily find out what he is looking for, and, even more, to perform this task in a handy and effective way [2]. Nowadays, a user with certain information requirements starting a navigational web session already finds tools able to make specific navigational aspects easier. For instance, there are some excellent browsers available to efficiently provide a great number of documents containing certain terms, or even to support searches of simple expressions based on these terms. Most of these browsers also permit to access the information included in the web through a thematic tree. However, the tools offered neither makes the user-system interaction easier, in a broad sense, nor do they solve the user's concrete problems, as they provide too generic solutions.

2. Web Interface Adaptivity

2.1. User profiling

Generally speaking, many HCI aspects have experienced much progress in recent years, but only if compared to other kinds of information systems. This is especially true in user interface adaptivity research [5], [10]. Obviously, when there is no adaptivity to user's skills or preferences, the interface designed for a wide group of users reduces the satisfaction degree for most of them. Consequently, to obtain more usable interfaces, it is necessary to know how these users are and in which way they behave in front of the interface. This is the reason why the so-called user profiles are established [1], [6]. These profiles allow grouping potential users in different sets by some common features. These features usually incorporate certain behavioral norms about the user interactions with the system. The inclusion of a user in a profile implies the adaptation of the application interface to this profile all the time, with the intention to make the interaction easier. But, the question is now if this general approach can be useful when applied to in learning systems.

When a student browses a didactic web, he tends to follow some repetitive guidelines according to the information requirements at any moment. It is not our purpose to establish the way in which these guidelines develop, neither to define them in a detailed way. Simply, we start from the idea that these guidelines exist, and basing our reasoning on a coherence principle, it may be concluded that there exists a certain connection between the student information requirements at a given moment and the information he has received along the session. This very valuable information has to be filtered to improve the future navigation.

Hence, it is not our desire at this moment to identify a specific student and to accumulate this information for use in the future, but rather to establish anonymous user profiles. This is non-persistent (or temporary) information used in the adaptation of the interface to the user requirements. The first point to establish is which kind of information about the learning session must be collected. Actually, all web browsers include mechanisms – help in the address bar, or link lists, for instance - for accessing the last recently visited pages in an easier way. They also offer support for the most frequently visited addresses by means of bookmarks. However, the intention of our research is to go a step further, making the access to recently visited contents easier. This is an ambitious starting point, because it is not easy to extract thematic information from the web documents as they are designed at this moment.

2.2. Semantic web and metadata

For this purpose, we need to engage into the new approach to the web called Semantic Web [12], [18]. This new web suggests enriching the information present in the traditional web with meta-information that describes more basic information [8]. In this sense, there are many projects in progress [16], [17], but nowadays most of the documents available on the web lack in describing metadata.

Title	Publisher	Format
Relation	Creator	Contributor
Identifier	Coverage	Subject
Date	Source	Rights
Description	Type	Language

Table 1. Dublin Core labels

Some projects define their own label set for metadata, being one of the most accepted the Dublin Core (DC) [4]. Dublin Core defines a set of 15 labels, although some people prefer a more extended set. The labels have names like *Title, Creator, Subject, Description*, etc. (see table 1). It is the *Subject* label that is of a special interest to our purpose. However, there are some questions that remain open. How can we interpret this field? Is this information useful by itself to our purposes?

2.3. Ontologies

Apparently, there has been little advance since some search sites approaches like Google [7] appeared in the market. The real relevant terms about the document content are present, but there is no possibility to interpret their sense. When a user browses a polysemic term in Google, the tool returns all the pages that contain this term in any of their senses.

On the other hand, the web browser forces a perfect lexical match, because it looks just for strings ignoring their semantic interpretation. However, the Semantic Web allows preventing this lack of semantic interpretation, by introducing the ontology concept [11], [13], [15]. An ontology establishes relations between all the terms that appear in the document descriptions.

An ontology like RDF [14] establishes a sentence format, <*subject*> <*verb*> <*predicate*>, in which every field is expressed in the sentence by means of the URL of the document that contains this element description. This allows, for example, defining term equivalences, as well as other kinds of relations. It is precisely this simplicity what makes RDF a very flexible ontology.

Actually, there are other ontologies, some of them designed for very different purposes in the Semantic Web. Some years ago, a large lexical database project was developed in the USA: WordNet [20]. WordNet is built on an ontology which is stricter than RDF. Its minimal information unit is the *synset*. A *synset* is a set of words – nouns, verbs, adjectives or adverbs - at the same grammatical category and with the same meaning in a given context. With these elements, and using a limited relational set, WordNet establishes a semantic lattice. Some of the default relations are synonymy, antonymy, hyponymy, meronymy, etc. (see Figure 1). This ontology provides enough expressiveness to help achieving our objectives and allows us to efficiently process the information.

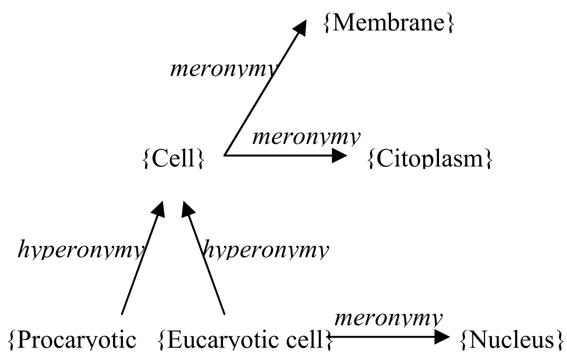

Figure 1. Main relations in WordNet

Now, starting from an ontology like WordNet, it is possible to associate pages where semantically related terms appear, that is to say, those pages that not necessarily represent the same lexical terms.

We will be able, for example in an organic chemistry context, to allow the user the access to pages about saturated hydrocarbons, where these pages do not really contain this term. Pages containing terms like paraffin or alcan that are equivalent to saturated hydrocarbon will now appear. We will also be able to discern a page about chaperons in its biochemical sense from other pages where it appears with other senses.

Now that the formalisms for our project have been established, the way they are being used will be explained in detail. It is our intention to apply the user interface adaptivity concept into the web application context at different levels. Up to date, there is a simple prototype available.

3. Framework description

3.1. Capturing the navigational aspects

When a student enters our web site, he can see the application window divided into two frames. On the left side, there is a smaller frame, where the suggestions are shown, while the rest of the window is the main frame, where the visited pages are loaded (see Figure 2).

A suggestion is simply a list of links to all documents in our site that are related, at a given degree, with the last ones the student has visited. We do not pretend at this moment to do global searches in Internet, but only to work in a concrete site context. This idea could be useful for academic web sites with a large number of scientific documents, where it is difficult to browse due to the amount and heterogeneity of the topics. Thus, when a student arrives to our web site, a cookie is created that stores all the themes which have appeared in the consulted documents as well as the frequency of appearance. When the student accesses a new document, the frequencies of appearance of all themes in the document that match his profile are refreshed, and the new themes are added to the profile. At any time, a list of documents matching his profile is shown to the user. This is performed by the server that looks for the metadata of our web site documents. We could have included DC labels into some documents, like, for example, in HTML pages through XML, but we have chosen a uniform treatment for all the possible documents. So every document must be accompanied with its description in a separate XML document. Thus, these are the files that are read in the server to obtain the list.

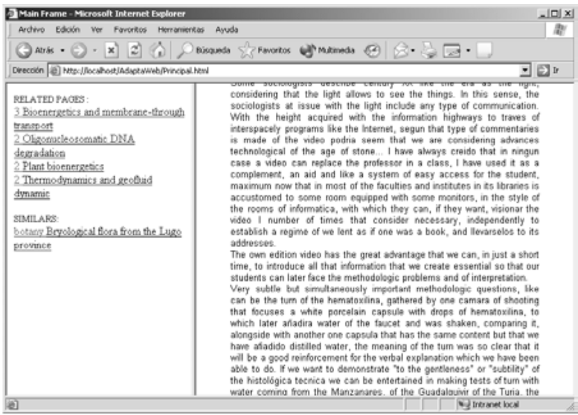

Figure 2. General layout of the web site

Notice, however, that we are not looking for a lexical equivalence, but rather for a certain semantic relation degree. To get it, there is an ontological information base, consisting of a data file set and its respective query function set. As you may observe in Figure 3, these files are divided into text blocks, each one containing the description of a *synset*. In the block header, we can find the *synset* terms, and, in a sequential way, a relationship list where the *synset* appears.

3.2. Managing the ontological information

Since a *synset* is enough structured, as we have seen, through an appropriate interface, we can obtain the semantic support that our model needs. On Table 2 all the interface methods are shown. In this table, we can see the methods defined to manage this ontological information. The first ones, through a convenient relationship, allow us the browsing of the semantic tree (remember Figure 1) and obtaining reachable terms by any length branches. Nonetheless, the last ones are really of a special interest to us. Let us focus on some of them.

```
0 @9169@ WORD_MEANING
  1 PART_OF_SPEECH "n"
  1 VARIANTS
    2 LITERAL "eucaryote cell"
      3 SENSE 1
      3 STATUS 99
  1 INTERNAL_LINKS
    2 RELATION "has_hyperonym"
      3 TARGET_CONCEPT
        4 PART_OF_SPEECH n
        4 LITERAL "cell"
          5 SENSE 1
    2 RELATION "has_mero_part"
      3 TARGET_CONCEPT
        4 PART_OF_SPEECH n
        4 LITERAL "membrane"
          5 SENSE 1
    2 RELATION "has_mero_part"
      3 TARGET_CONCEPT
        4 PART_OF_SPEECH n
        4 LITERAL "nucleus"
          5 SENSE 1
  1 EQ_LINKS
    2 EQ_RELATION "eq_synonym"
      3 TARGET_ILI @27535@
```

Figure 3. A file containing ontological information

- **Closeness**: quantifies the semantic closeness of two terms. This method calculates the distance between the tree nodes and the relevance of the relationships used. It is invoked to choose the best tailored sense in a context (where this function gets the highest value).

- **Sense_of**: establishes the sense of a term within a context (a list of unknown sense terms). Thus, this function allows jumping from the lexical unit used at metadata level to the semantic unit needed to access the ontology. This step is necessary if we want the phases to be independent. Thus, when a new document is added, and its subjects are written, it is not necessary to know the ontology. The temporal information stored about the session is also semantic information. Then, to match new pages into the profile, we must make a conversion.

There is also a term closeness scale to be introduced. We define closeness degree 0 when two terms match lexically. Degree 1 is used if the terms are synonyms, degree 2 if they are joined by one step of the first relationships, and so on.

Function	Returned Value
Ret_dir_rel	Set of terms related to a given one through a relation in a single step
Ret_relationed	Set of terms related to a given one through a relation in any number of steps
Return_relationed	Starting from a term and its sense, all terms related through a list and degree
Is_relationed_s	Are two terms related through a relation (in any number of steps)?
Has_common_ ancestor	Starting from two terms, their senses and a relation, it determines if they have a common ancestor
Synset	Set if terms of a *synset*
Sense_set	Set of senses of a term
Unique_sense	Is a term monosemic?
Closeness	Closeness function
Sense_of	Sense of a term from a context

Table 2. Ontological information query functions

This mechanism allows configuring the web site to work at different precisions, because the respective functions have degree parameters.

This interface allows using our ontology to interpret document metadata and to extract their semantic information. Sometimes, an isolated metadata can still have semantic ambiguity, but then, our ontology allows solving it through the context.

3.3. Exploiting the information

As told before, the server is the one that can explore the web site document features. We are interested in obtaining the descriptive theme set for each document. Starting from this term list, and through the tools studied, it is possible to quantify the user's interest in the document. We could say that we have software agent aided navigation [19].

From this perspective, an automaton maintains an updated list of links to help the user to reach his objectives in a handy and efficient way. When a page matches some user interest themes, this page is proposed in the left frame. We must give priority to most tailored links, even showing only the best ones.

We included some mechanisms to quantify page closeness to all the user consulted themes, but perhaps, these themes are not equally relevant. In a way, we must be able to study user behavior, because all the stored information is not equally useful at a given moment.

On the other hand, during a same session a user may consult on different thematic areas. Therefore, different behavioral patterns are created in a same navigation session, getting different areas of the global ontology. This, and other factors, must be studied in detail for the purpose of tailoring possible user interest evolutions.

4. Conclusions

At the beginning of this paper, we considered the application of user interface adaptivity principles to the didactic web application area. We affirmed that this was not an easy task due to the great number of features involved in this kind of applications. Thus, when we began to analyze the way in which our prototype puts our ideas into practice, we assumed certain restrictions. These restrictions affect the number of accessible documents (we work just in a controlled web site), and the complexity of the user profile (we only store a limited set of visited themes), and allow us to aid the user through an intelligent navigational session by means of an adaptive interface.

But there is still much work to do in our attempt to provide the user with really useful information through a guided interface. The analysis carried out on the first prototype has allowed, as expected, to study the features incorporated to the user model, the limitations, and some hints to improve our initial intentions.

Acknowledgement

This work is supported in part by the Spanish Junta de Comunidades de Castilla-La Mancha PBC-03-003 grant.

References

[1] Chan, P. A non-invasive learning approach to building web user profiles. Proceedings of the KDD-99 Workshop on Web Analysis and User Profiling, pp. 7-12, 1999.

[2] Cueva, J.M., González, B.M., Aguilar, L., Labra, J.E. and del Puerto, M. (Editors). Proceedings of the International Conference on Web Engineering, ICWE 2003, Lecture Notes in Computer Science 2722, 2003.

[3] Dix, A., Finlay, J., Abowd, G. and Beale R. Human-Computer Interaction. Prentice-Hall, 1998.

[4] Dublin Core Metadata Initiative.
http://www.dublincore.org

[5] Fernández-Caballero, A., López-Jaquero, V., Montero F. and González, P. Adaptive interaction multi-agent systems in e-learning / e-teaching on the web. Third International Conference on Web Engineering, ICWE 2003, Lecture Notes in Computer Science 2722, pp. 144-153, Springer-Verlag.

[6] Fons, J., García, F. J., Pelechano, V. and Pastor, O. User profiling capabilities in OOWS. Proceedings of the International Conference on Web Engineering, ICWE 2003, Lecture Notes in Computer Science 2722, pp. 486-496, 2003

[7] Google. http://www.google.com

[8] Klamma, R., Hollender, E., Jarke, M., Moog, P. and Wulf, V. Vigils in a wilderness of knowledge: metadata in learning environments. Proceedings of IEEE International Conference on Advanced Learning Technologies, ICALT 2002, pp. 519 -524, 2002. IEEE Computer Society, Los Alamitos, CA.

[9] López-Jaquero, V., Montero, F., Molina, J.P., Fernández-Caballero, A. and González, P. Model based design of adaptive user interfaces through connectors. 10th Workshop on Design, Specification and Verification of Interactive Systems, DSV-IS 2003, Lecture Notes in Computer Science 2844, pp. 245-257, 2003.

[10] López-Jaquero, V., Montero, F., Fernández-Caballero, A. and Lozano, M.D. Towards adaptive user interface generation: One step closer to people. Proceedings of the 5th International Conference on Enterprise Information Systems, pp. 97-103, 2003.

[11] Macías, J. A. and Castells, P. Tailoring dynamic ontology-driven web documents by demonstration. Proceedings of the 6th International Conference on Information Visualisation - International Symposium of Visualisation of the Semantic Web, page 535. IEEE Computer Society, 2002.

[12] Olsina, L., de los Angeles M., Fons, J., Mara S. And Pastor, O. Towards the design of a metrics cataloging system by exploiting conceptual and semantic web approaches. Proceedings of the International Conference on Web Engineering, ICWE 2003, Lecture Notes in Computer Science 2722, pp. 324-333, 2003.

[13] On-To-Knowledge: Content-driven Knowledge-Management through Evolving Ontologies. http://www.ontoknowledge.org

[14] Resource Description Framework (RDF).
http://www.w3.org/RDF

[15] Tamura, Y. Domain-oriented approach to the reuse of learning knowledge: An overview. Proceedings of IEEE International Conference on Advanced Learning Technologies, ICALT 2002. IEEE Computer Society, Los Alamitos, CA.

[16] The Maryland Information and Network Dynamics Lab Semantic Web Agents Project. http://www.mindswap.org/

[17] The Semantic Web Community Portal, Markup Language and Ontologies.
http://www.semanticweb.org/knowmarkup.html

[18] Wahlster, W., Lieberman, H. and Hendler, J.A., Spinning the semantic web: Bringing the world wide web to its full potential. The MIT Press 2003.

[19] Weiss, G. (Editor). Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. The MIT Press, 2000.

[20] WordNet: A Lexical Database for the English Language. http://www.cogsci.princeton.edu/~wn