Mathematica Balkanica

New Series Vol. 25, 2011, Fasc. 5

Semantic-Oriented Architectures and Use of Ontology for Organizing Adaptive Search in e-Learning Environments

Albena Turnina

In this study brief review of adaptive search and Semantic Web technologies is outlined as some of the basic concepts in the field are presented and one particular solution is shown as well.

MSC2010: 97C30, 97C70

Key Words: Adaptive hypermedia, WEB systems, e-learning

In recent years the process of increasing of available information on the Web can be observed which is complemented by increase in informational needs of users. This necessitates the creation of models that facilitate users in access to information tailored to his/her specific characteristics and interests. These models can be used in two type environments – in Adaptive Hypermedia systems, and in Web personalization systems. Web personalization is a process of adapting content and structure of web site according to specific needs of particular user using navigation history of his/her surfing. Adaptive hypermedia systems implement adaptation by manipulating structure of links or by changing presentation of information based on explicitly specified user profile (Eklund & Sinclair, 2000; De Bra et al., 1999; Brusilovsky, 2001).

There exist similarities and differences between the Web personalization systems and the Adaptive Hypermedia systems, namely: both systems are based on modeling the user profile, which subsequently serves as a basis for making further recommendations. The differences come from the different ways in which both systems create that profile, whereas the Web personalization systems use tracking algorithms to follow user in his/her surfing on the Web, the Adaptive hypermedia systems build the user profile based on user interaction with the system (Mobasher et al., 2002).

According to (Cingil et al., 2000) steps needed to be performed in order to create Web personalization are as follows:

- to collect data from web
- to model and categorize that data
- to analyze collected data and to determine the appropriate action.

Essential step for achieving Web personalization is acquirement of explicitly provided user data and according to that criterion the Web sites can be differentiated as follows:

- web sites that require registration only
- web sites that require registration and more explicit set of user information describing his/her interests and characteristics.

There are different types of personalization that (Tsianos et al., 2009) distinguished as follows: link personalization, content personalization, context personalization, authorized personalization and humanized personalization.

The necessity for personalization of the content presented to the end user is recognized a long time ago –in the early 90 and since then have been developed and presented many systems, giving its particular answer to the question how to build effective personalization systems. The need of personalization over time increases along with the increased volume of resources distributed along the network, as well as with the growth in number of users using information resources. The change of proportion of younger against older consumers over time can be observed as statistics shows that the number of older users over time increases. The adaptive systems leverage the differences amongst the users surfing along the Web as they adapt presented information in accordance with particular user profile.

Adaptive systems allow each particular user to be characterized by a number of personal characteristics, including not only demographic characteristics as gender and age but also another attributes as level of knowledge and experience in particular field, cognitive abilities, visual characteristics etc. The users characteristics can be described and classified according to the user model and subsequently can allow the clustering of users on the basis of the similarities of their profiles.

(Tsianos et al., 2009) propose the following paradigms providing personalization:

- Content-based filtering - encompass systems that track user browsing behavior and recommend topics that are similar to items previously liked by the same user. This technique has some shortcomings, for example so called "over-specialization". Content-based filtering systems are based on characteristics of user profile and their recommendations relate solely to the items to which users have already shown interests. (Shahabi &

Chen, 2003; Mobasher et al., 2002) identified other shortcomings of this type of systems, including the lack of semantic relations between objects. Example systems, content-based filtering based on are WebWatcher and Letizia.

- Collaborative filtering systems that use cumulative experience of large groups of users in order to facilitate finding of a particular topic from a user. This is achieved by recommending the user to see a content already seen and liked by other users with similar profiles. This approach is based on the assumption that users with similar behaviors have similar interests. Employed algorithms investigate similarities between users and recommendations are given according to the preferences of most similar neighbor. Example systems based on collaborative filtering are Yahoo, Excite, Microsoft Network, Net Perceptions, Amazon.com, and CDNOW.
- Demographic-Based Filtering covers systems that use demographic similarities between users as a basis for their grouping. Recommendations are given by the system with accordance to age, gender and other demographic characteristics of the users (Pazzani, 2005).
- Rule-Based Filtering in these systems users are encouraged to respond to a predefined set of questions. The next step the system takes is to recommend products and services deemed appropriate for that user on the base of provided answers. Examples of such systems include Dell, Apple Computer, and Broadvision.
- Agent Technologies these kind of systems operate on the base of agents that perform tasks on behalf of the user with a degree of autonomy that allows them to be regarded as personal assistants (Delicato et al. 2001). Example of systems using agent technologies are ARCHIMIDES, Proteus, WBI, BASAR, eRACE, mPersona, Fenix ??system and SmartClient.
- Cluster Models this type of technology is found in e-commerce and is based on the division of consumers in clusters, as each user is classified into one of the existing clusters. Amazon.com is based on this kind of model.
- Hybrid systems combine two or more technologies to overcome the draw-backs of each of them (Balabanovic, 1998).

(Oard, 1997) developed a model that describes the Content-based filtering as consisted of following methods considered essential for its functioning:

- the method for presentation of documents within the domain
- the method for consumer information needs

- the method for making comparisons between the two.

A popular and largely used model for document presentation and user profile presentation in the Content-based filtering system is the vector space model. In this model any document or profile is presented as a set of keywords or as n-dimensional vector. This presentation serves as a basis for comparison between the document and profile and allows further measurement of the degree of similarity between both. If similarity between document and profile falls in predefined threshold then the document is considered relevant to user.

A largely known method for calculation of probability for document relevance to user is the Bayesian probabilistic classifier. The popularity of the vector space model over other methods and models is mainly due to its simplicity and proven effectiveness. The vector space model has some limitations due to the potential loss of information as the document is separated into words without taking into account the meaning presented by sentences or phrases. These shortcomings can be overcome by other techniques and technologies taken from the fields of Artificial Intelligence and Information Retrieval.

According to (Abidi, 2009) the Intelligent Personalization is dynamic and intelligent adaptation which is achieved on the base of information found in a user profile. An example of an intelligent system incorporating intelligent personalization is AdWISE (Adaptive Web-based Information and Services Environment), which uses methods and techniques taken from the field of Information Retrieval and Artificial Intelligence.

(Eklund & Sinclair, 2000; Brusilovsky 2001) argues that the adaptivity is a specific feature that facilitates the navigational difficulties of users by means of user profiles encompasses specific users characteristics.

Adaptive Hypermedia systems apply adaptivity by manipulating the structure of links or by changing the way information is presented based on the particularities of the user, expressed in explicit profile. (Eklund & Sinclair, 2000; De Bra et al., 1999; Brusilovsky, 2001).

(Brusilovsky, 2001) a classification of user's characteristics, used in Active Hypermedia systems, is as follows:

- user's knowledge
- user's purpose
- user's background
- user's experience with hypertext
- user's interests
- environment

In accordance to (Smyth, 2009) the techniques for navigation and search can be characterized as reactive modes of data access because they need the explicit action of the user and thus differ from the systems that make recommendations, which act in proactive mode. At the moment, the most popular search is the keyword search and the most efficient and largely used search engines are based on the keyword search because of the great simplicity and efficiency of this method. Despite its efficiency the keyword search method has some shortcomings, namely high recall (too many returned results) and low precision (insufficient relevance). The presented outcome from the user's search often includes documents with low relevance to query or duplication of search results or missed relevant documents because of the inadequate consideration of synonyms. All of these shortcomings can be overcome by use of semantic search techniques and technologies which allows contextualization of query and relevance of the presented results to the specific user profile. The adaptive search provides the following abilities which increase the precision of returned to user query results, namely:

- ability to take into account the context of query
- ability to formulate a natural language query
- ability to adapt interface according to user preferences
- ability to provide additional resources and recommendations

In order to effectively implement the adaptive search techniques creation of user profile is needed. There are different approaches and methods for creating such profile - in one hand the needed information can be provided explicitly by the user, on the other hand the data can be collected by a system tracking user behavior and activities. The user profile can be static in this case the profile do not change over time or it can be designed to be dynamic, allowing the changes in user behavior to be accurately recorded and implemented.

The popular and easy method for profile creation is on the base of inputted keywords, which serves as a representation of user's interests and informational needs. The profiles created by means of the semantic technologies can present significant advantages because of the added meaning to data which allow further reasoning over collected facts. (Dolog, 2005) proposed a model of the learner based on semantic technologies.

In recent years the emergence of new technologies of the Semantic Web that extend the capabilities of Web by adding semantic information to the resources and services are increasingly looking for effective ways to create personalization and adaptivity based solely on them. Among the advantages of this approach is the ability for formal representation of knowledge in machine understandable form and consequent machine inference over that knowledge. The technologies of the Semantic Web and more specifically ontology are effective

means for expressing user knowledge as well as domain knowledge as hierarchy of concepts and their relations. Ontology represents a substantial part of the Semantic Web technologies which allow granular and highly expressive knowledge of domain to be represented.

One way to implement personalization is to develop user profile ontology as well as domain ontology in order to be able to model characteristic of user and develop recommendation system on the base of the profile. This approach is implemented in this particular study in which we propose a system based on two ontologies – one for the domain of interest and another for the user profile. The user profile ontology up to now is under development and will be presented in the future work so in this study we will show the results achieved so far.

In the next section a brief overview of semantic technologies will be provided. The term Semantic Web was coined by Sir Tim Berners-Lee to encompass framework of standards and technologies used for semantic markup and annotation of data on the Web. Semantic Web is seen as the next generation in the development of the Web that will allow easier transmission, search, mixing and processing of data and make them understandable by machines. The structure of the Semantic Web can be described by a seven layer's model proposed by Tim Berners-Lee. In this model, on the bottom layer are located Unicode character sets and Uniform Resource Identifiers (URIs). The former makes possible use of all world alphabets whereas URIs, which is an extension of the popular URLs, serves to provide a unique representation of elements by strings of symbols identifying Internet resources. At the next layer to the up is presented the eXtensible Markup Language (XML), which sets the basic syntax for describing data. Above XML is located the Resource Description-Framework (RDF), which provides semantic structure based on the use of triples "subject, object, predicate." Within the same layer, the RDF Schema (RDFS) allows the definition of structures such as classes and subclasses. Above them in the layer's model are the available ontological vocabularies which expand the logical structure of RDFS. The so-called Logical layer operations are performed on the claims defined in ontological language. Proof layer provides evidence as a necessary condition for commercialization. The top layer includes rules governing Trust. The top three layers are under development and not yet entered the stage of Recommendation (Recommendation) by 3WC.

Ontology as an essential technology for delivery of semantics is of great significance in the Semantic Web. The term ontology comes from ancient Greek: ontos means life, existence, and logos means science, study, theory or term ontology stands for science of life. In the field of the Artificial Intelligence the term ontology is used to describe formally represented knowledge based on a conceptualization. According to the famous definition of Gruber (Gruber, 1993) ontology is "a formal explicit specification of shared conceptualization." Conceptualization involves the description of the set of concepts, their relations and facts and rules that governs them. Ontology is defined as data model

that describes the "things that exist" in some particular domain (concepts, their properties, attributes, facts, rules and relations) in a coherent and formal way. Ontology can be regarded as a standardized model that provides basis for common understanding of a field and this model can be shared between people and / or computing systems. The set of statements constituting the ontology use First order logic in order to be able to reason about them and draw new facts and derive conclusions. In its simplest form the ontology is a hierarchy of concepts (classes) known as taxonomy. In more complex cases the appropriate axioms can be used to further express relations between concepts and to limit interpretations of their meaning. In this sense,the ontology is a knowledge base, describing facts which are always valid within a given domain. Ontology can be created automatically by means of special techniques taken from the field of machine learning, statistical analysis and NLP or can be constructed by hand. Each of two methods poses some advantages and disadvantages.

According to (Middleton et al., 2002) ontology-based representations are richer, more accurate and less ambiguous method for describing knowledge as a whole and particularly the user profile than traditional approach based on keywords. Innovative methods for processing data include the ability to automatically extract or learn domain ontology and the hierarchy of classes based on the data of the domain. In some cases hand construction of ontology is preferable approach because although it is more laborious it allows full control over the created ontology and also provides the possibility for future interoperability.

A main ontology component is a concept or class, which is a description of the nature of an object. Concepts are characterized by associated properties that link individuals of different concepts as well as link individuals to their values. There are three types of properties: Object, Datatype and Annotation properties, the last add metadata to classes, individuals and object/data type properties. Properties set a hierarchical (is-a relations) and can have some restrictions. The rules in the ontology are defined by axioms and constraints that are always true and are useful for verifying the correctness of input data. Currently, there are many classifications of ontology, according to Fancel (Fensel et al, 2004) ontologies can be classified as follows: general ontologies, application ontologies, task ontologies, representational ontologies, etc.

The ontology presented in this study is named "Student international class" and is developed in order to describe domain of student class with international students belonging to five different nationalities. The used method for ontology construction is by hand. The used tools are an ontology editor Protege, chosen because of its good features and because it is distributed free; reasoner Pellet used to validate ontology; viewer OntoViz. The ontology language is OWL DL because of its ability for automatic reasoning based on description logic.

Description of ontology:

- 1. Five different countries
- 2. Students have to be divided in three distinct classes based on age, gender and nationality and such that in each class have to be enrolled at least one girl and at least one Nigerian student.
- 3. Each student class must have at least ten students
- 4. All students must be fluent in English
- 5. National holidays of all student will be declared as free days
- 6. Food restriction must be taken into account in order to avoid restricted food in lunch menu

Ontology classes, properties, restrictions and individuals are as follows:

I. Classes

1. Taxonomy

- - Class Person with subclasses Girl, Boy, Student, Teacher
- - Class Nationality with subclasses French, English, Italian, Nigerian, Indian
- - Class Classroom
- - Class Subject
- - Class with subclasses LowerAge, MiddleAge, HighAge
- - Class Language
- - Class Lunch
- - Class Food with subclasses Salad, Soup, Dessert, Dish with subclasses Vegetarian, Unvegetarian
- - Class Country
- - Class Holiday with subclasses (EnglishHolidais, FrenchHolidays, ItalianHolidays, NigerianHolidays, IndianHolidays)
- Disjoint classes individual which has been asserted to be a member of one of the classes in the group cannot be a member of any other classes in that group
 - -Student Teacher
 - -Boy Girl

- -Vegetarian Unvegetarian
- - LowerAge HighAge MiddleAge
- -Salad Soup Dessert Dish
- - Italian Holidays - Indian Holidays - Englisg Holidays - French Holidays
 - NigerianHolidays
- - Italian English Indian Nigerian French

3. Superclasses

- -class Person superclass of classes Boy and Girl (covering axiom)
- -class Person superclass of classes Student and Teacher
- -class Dish superclass of classes Salad, Soup, Dish and Dessert

4. Example class realization

II. Properties - represents relationships between two individuals. There are three types of properties: Object, Datatype and Annotation properties

- 1. Object properties link an individual to an individual
 - -has_nationality Functional property
 - -nationalitiOf Inverse functional property of has_nationality
 - -has_holiday Inverse of holidayOf

- -holidavOf
- -has_restrictedFood Inverse of restrictedFoodOf
- -restrictedFoodOf
- -has_language Inverse of languageOf
- -languageOf
- -has_teacher Inverse of teacherOf
- -teacherOf
- -has_student Inverse of studentOf
- -studentOf
- -has_classmate Symetric property
- -has_lovelyFood Inverse of lovelyFoodOflovelyFoodOf
- -comprises_student has_classroom Inverse of classroomOf
- -classroomOf
- -has_birthCountry Functional property
- -birthCountryOf Inverse Functional of has_birthCountry
- -has_subject
- -has_launch
- 2. Data type Properties link an individual to an XML Schema Datatype value or an rdf literal
 - -has_databirth
 - -has_age
- 3. Annotation properties add metadata to classes, individuals and object/data-type properties
- 4. Example properties realization

III. Restrictions are used to restrict the individuals that belong to a class

- Quantifier Restrictions existential, universal
- Cardinality Restrictions
- hasValue Restrictions
- 1. Existential Restrictions for a set of individuals, an existential restriction specifies the existence of a (i.e. at least one) relationship along a given property to an individual that is a member of a specific class

1.1. Necessary

- Student has_holliday some Hollyday
- Student has_lovelyFood some Salad
- Student has_lovelyFood some Soup
- Student has_lovelyFood some Dish
- Student has_lovelyFood some Dessert
- Student has_restrictedFood some Food
- Student has_teacher some Teacher
- Student has_classmate some Student
- Student has_launch some ((Salad and Dish and Dessert) or (Soup and Dish and Dessert))
- Classroom: has_student some Nigerian
- Classroom: has_student some Girl

1.2. Necessary and Sufficient

- Student has_language has EnglishLanguage
- 2. Universal Restrictions the only relationships for the given property that can exist must be to individuals that are members of the specified class
 - Student has_nationallity only Nationallity
 - Student has_birthCountry only Country,
 - Class has_teacher only Teacher

3. Cardinal

- Student has_nationallity exactly 1
- Class has_teacher min 1
- Class comprises_student min 10

- Class has_classroom exactly 1
- HighAge comprises_student min 10
- LowerAge comprises_student min 10
- MiddleAge comprises_student min 10

4. Datatype

- Student has_age has Integer
- LowAge has_age max 10
- MiddleAge has_age min 10
- MiddleAge has_age max 14
- HighAge has_age min 14
- HighAge has_age Max 16
- 5. Example restriction realization

has_launch some ((Salad and Dish and Dessert) or (Soup and Dish and Dessert))

```
<owl:Class>
           <owl:intersectionOf rdf:parseType="Collection">
             <owl:Class rdf:about="#Salad"/>
             <owl:Class rdf:about="#Dish"/>
          <owl:Class rdf:about="#Dessert"/>
           </owl:intersectionOf>
         </owl:Class>
         <owl:Class>
           <owl:intersectionOf rdf:parseType="Collection">
             <owl:Class rdf:about="#Soup"/>
             <owl:Class rdf:about="#Dish"/>
             <owl:Class rdf:about="#Dessert"/>
           </owl:intersectionOf>
        </owl:Class>
      </owl:unionOf>
   </owl:Class>
 </owl:someValuesFrom>
</owl:Restriction>
```

IV. Individuals

- Instance of Student: Mo, Alan, Peter, July, Iv, Din, Pin, Ho, Su, Jane, Kalan, Marko, Mariet, Suzan, Evelyn, Mishel, Irin, Van, Cin, El, Ana, Maja, Mariam, Estefan, Dan, Lily, Jerom, Fil, Fani, Za, Magombe
- Instance of Teacher: miss Elan, miss Violet, miss Jin
- Instance of Class HighAge ClassA, LowerAge ClassB, MiddleAge ClassC
- Instance of Food Soup: Borsh, OnionSoup
- Instance of Food Salad: Cesar, Shopska
- Instance of Food Dish: Chicken, Pork, Fish, GrillVegetables
- Instance of Food Dessert: Brule, IceCream
- Instance of Country Italy, Niger, India, England, France
- Instance of Language EnglishLanguage, ItalianLanguage, FrenchLanguage, IndianLanguage, NigerianLanguage
- Instance of Classroom Classroom1, Classroom2, Classroom3

- Instance of Subject - Math, Geography

V. Query with SPARQL

Most forms of SPARQL query contain a set of triple patterns called a basic graph pattern. Triple patterns are like RDF triples except that each of the subject, predicate and object may be a variable. A basic graph pattern matches a subgraph of the RDF data when RDF terms from that subgraph may be substituted for the variables and the result is an RDF graph equivalent to the subgraph.

1. Use case one:

We want to know the all national holidays for different nationality students in order to announce these days as non learning days.

```
SELECT ?holiday
FROM <http://www.owl-ontologies.com/Ontology1298459866.owl>
WHERE {
?holiday
    <http://www.owl-ontologies.com/Ontology1298459866.owl#holidayOf>
?student }
```

2. Use case two:

We want to know the names of all students which have the concrete teacher, for example Miss Violet

```
SELECT ?student
FROM <a href="http://www.owl-ontologies.com/Ontology1298459866.owl">http://www.owl-ontologies.com/Ontology1298459866.owl</a>
<a href="http://www.owl-ontologies.com/Ontology1298459866.owl#missViolet">http://www.owl-ontologies.com/Ontology1298459866.owl#missViolet</a>
<a href="http://www.owl-ontologies.com/Ontology1298459866.owl#missViolet">http://www.owl-ontologies.com/Ontology1298459866.owl#missViolet</a>
```

3. Use case three:

We want to know all students which are vegetarian

The next step will be development of user profile ontology encompassing characteristics of user profile and accounting how the characteristics affect the returned in the result of the query results. Among the advantages of using ontology to model profile is considered the context associated with the interests of each customer. To achieve that objective we can use different approaches more specifically automatic construction of ontology or construction by hand. The pros and cons for both methods will be evaluated. (Pretschner and Gauch, 2004) use navigation history of user to create a user profile ontology and take into account the time spent by the user on some particular document. The advantage of this approach is the ability to constantly update the profile concepts in accordance with the future activity of the user. An example of running and widely popular ontology, which serves to describe people's interests and relationships is FOAF - Friend of a friend ontology. Using ready ontology allows easy integration of heterogeneous sources and high degree of interoperability.

The presented in this study ontology has been tested and validated with reasoner – Pellet and the obtained results shows that the ontology does not have logical errors. The chosen language OWL DL allows machine inference. The ontology is queried by means of semantic query language SPARQL and the results fully justify the expectations.

References

- [1] Abidi S., (2009) Intelligent Information Personalization: From Issues to Strategies. Intelligent User Interfaces, pp 118–147.
- [2] Balabanovic, M. (1998). Learning to Surf: Multiagent Systems for Adaptive Web Page Recommendation. PhD thesis, Department of Computer Science, Stanford University
- [3] Brusilovsky P., Adaptive Hypermedia. (2001). User Modeling and User-Adapted Interaction, 11: 87-110
- [4] Cingil I., Dogac A., and Azgin A. (2000). A broader approach to personalization, Communications of the ACM, Vol. 43, No. 8.
- [5] De Bra P., Brusilovsky P., and Houben G. (1999). Adaptive Hypermedia: From systems to framework, ACM Computing Surveys 31(4).
- [6] Delicato F., Pirmez L., and Carmo L. (2001). Fenixpersonalized information filtering system for WWW pages, Internet Research: Electronic Networking Applications and Policy, Vol. 11, No. 1, pp. 42-48
- [7] Dolog, P. and Schafer, M. (2005). Learner modelling on the semantic web. In PerSWeb05 Workshop on Personalization on the Semantic Web in conjunction with UM05

[8] Eklund J. and Sinclair K. (2000). An empirical appraisal of the effectiveness of adaptive interfaces of instructional systems. Educational Technology and Society 3 (4), ISSN 1436-4522.

- [9] Fensel D., Sycara K., Mylopoulos J., (Editors) (2004). The Semantic Web. ISWC 2003, Springer-Verlag.
- [10] Gruber, T (1993). A translation approach to portable ontologies. Knowledge Acquisition, 5 (2), pp. 199-220.
- [11] Middleton, S., Alani, H., Shadbolt, N., and Roure, D. D. (2002). Exploiting synergy between ontologies and recommender systems. In Semantic Web Workshop.
- [12] Mobasher, B. (2005). Web usage mining and Personalization. In Munindar P. Singh (ed.) The Practical Handbook of Internet Computing. CRC Press.
- [13] Mobasher B., DaiH., Luo T., Nakagawa M., and Wiltshire J. (2002). Discovery of aggregate usage profiles for Web personalization. Data Mining and Knowledge Discovery, Vol. 6 (1), pp. 61–82.
- [14] Oard, D. (1997). The state of the art in text filtering. User Modeling and User-Adapted Interaction
- [15] Pazzani J. M. (2005). A framework for collaborative, content-based and demographic filtering. Artificial Intelligence Review, December 1999, vol. 13, no. 5-6, pp. 393-408 (16).
- [16] Shahabi C. and Chen Y. (2003). Web Information Personalization: Challenges and Approaches, In the 3nd International Workshop on Databases in Networked Information Systems (DNIS 2003), Aizu-Wakamatsu, Japan.
- [17] Smyth B., (2009) Case Studies in Adaptive Information Access: Navigation, Search, and Recommendation. Intelligent User Interfaces pp.36–60.
- [18] Tsianos N., Germanakos P., Lekkas Z., Mourlas C. and Samaras G. (2009). An Assessment of Human Factors in Adaptive Hypermedia Environments Intelligent User Interfaces, 2009, pp 1–35.

Sofia University
"St. Kliment Ohridski"

e-mail: turnina@fmi.uni-sofia.bg