

Designing a Personalized Semantic Web Browser

Melike Şah, Wendy Hall and David C De Roure

IAM Group, School of Electronics and Computer Science, University of Southampton
{ms305r, wh, dder}@ecs.soton.ac.uk

Abstract. Web browsing is a complex activity and in general, users are not guided during browsing. Our hypothesis is that by using Semantic Web technologies and personalization methods, browsing can be supported better. However, existing personalization mechanisms on the Web are obstructive; users need to log in to multiple websites and enter their personal information and preferences, and the profiles are different for each site. There is a need for generic user profiles, which can also support the user's browsing. In this paper, we propose a novel Semantic Web browser using an ontology-driven user modeling architecture to enable semantic and adaptive links. We also introduce a new behavior-based user model. With our approach, users need to log in to their Web browser only and personalization is achieved on different websites.

Keywords: Semantic Web Browser, Semantic Web, User Modeling, Ontology, Personalization

1 Introduction

Searching and browsing are two important information filtering activities on the Web. Usually, users use search engines for finding Web resources but this is only half of the story. When users follow a link from search results, they have to read and understand page content and in general they are not guided during browsing. Browsing is a complex activity and its nature is not understood well. According to Bawden [1], the activity of browsing can be categorized into three groups: purposive browsing (looking for a definite piece of information), capricious browsing (randomly examining material without a defined goal) and exploratory browsing (deliberately searching for inspiration). Cove and Walsh [2] also divide browsing into three categories: search browsing (searching for defined information), general purpose browsing (looking for items of interest) and serendipity browsing (random). Based on these definitions, we can say that browsing tends to be used in three broad senses: a purposeful activity (directed), searching for inspiration (semi-directed) and capricious behavior (undirected). In our opinion, user profiles should contain such information.

Nowadays, personalization is supported by many websites on the Web (e.g. Amazon, Google). However, they require users to log in to multiple websites and the user profiles change from site to site. There is a need for generic user profiles and personalization architectures, which can achieve adaptive hypermedia on diverse websites. Our hypothesis is that Semantic Web technologies can offer the solution to

these problems. Semantic metadata can be used for finding related information during browsing. Additionally, ontology-based user profiles are interoperable, and they can be easily extended and combined with semantic metadata on the Web.

COHSE [3] and Magpie [4] are two semantic-enabled systems that aim to provide useful browsing hyperlinks using semantic metadata. However, they paid little attention to the user's role and they do not supply adaptive links or contents. In addition, they both use back-end databases for semantic linking.

In user modeling, there have been several studies. IMS LIP and IEEE PAPI are well known user modeling standards. Although these standards can be applied to any domain, they mainly developed for learners in educational hypermedia. The Ontologging project [5] develops a user ontology in the context of knowledge management systems. However, the user model is specific to a certain domain. In [6], a general meta-ontology is developed for modeling user and adaptive hypermedia methods.

In this paper, we propose a novel personalized Semantic Web browser architecture, which we named SemWeB. SemWeB is able to annotate Web pages with semantic metadata using ontologies and provides adaptive and semantic hyperlinks on different websites. For semantic linking, we use Web as source for linking (no back-end database is used). A new behavior-based and an ontology-driven user modeling architecture is also integrated into SemWeB and user profiles can be implicitly and explicitly updated with semantic metadata. This paper presents our user ontology, user modeling approach and ongoing work on the Semantic Web browser.

2 Architecture of SemWeB

SemWeB is a browser extension of the Mozilla Firefox Web browser (Fig. 1). SemWeB extends the Web browser with a vertical sidebar. The sidebar has two tabs: the navigation tab and the personalization tab. The navigation tab is used for highlighting ontological concepts found on the page and adding semantic and adaptive links. The personalization tab is used for updating user profiles. SemWeB annotates Web pages using ontologies and an ontology-driven lexicon based on a modified GATE framework [7]. GATE is a general text engineering architecture for extracting named entities from text. We are using rules and gazetteers for IE. Also, we extend GATE with a lookup service and annotation storage unit. Lookup service returns the URIs of found concept instances and annotation storage unit creates and stores semantic annotations as XML files at server-side. Because IE requires some pre-processing (creating lexicons, etc.), SemWeB uses predefined ontologies, particularly ECS ontology [8]. SemWeB can also be adapted to different ontologies.

Semantic links are inserted, when the user highlights a concept from the navigation tab. SemWeB embeds icons next to recognized instances on the Web page. Then, user is required to click the icon. Once user clicks, URI of the instance is sent to the server. First, server dereferences URI using HTTP content negotiation and finds possible link anchors and targets. More URI dereferencing performed and related URIs on the Web is searched. Finally, semantic links are presented in a new Web page at the browser.

To benefit personalization, users need to register and log in to SemWeB from the personalization tab. User profiles are kept at a server-side triple store. Additionally,

we created a profile editor. By using this editor, users can explicitly add, remove, change information to their profiles from the personalization tab. As well as, users can be explicitly assigned to expertise, interests and goals from the interface of SemWeB. Implicit user characteristics, such as browsing level and browsing type are automatically updated by SemWeB.

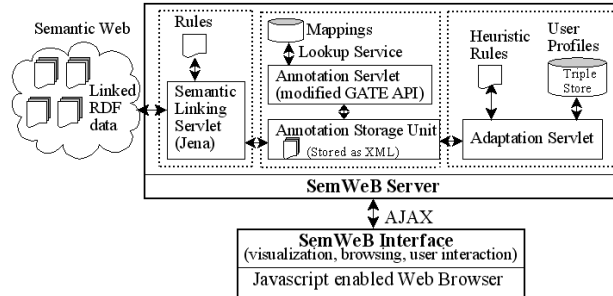


Fig. 1. The architecture of SemWeB

3 The Proposed User Model

In user modeling, IMS LIP and IEEE PAPI are well known standards. Although these standards can be applied to any domain, they do not contain information about browsing behaviors. Thus, we developed a new behavior-based user model, which can also be applied to different domains. In our model, currently we use seven categories: identification, preference, security, *browsing goal*, *interest*, *expertise* and *browsing behavior* (our main contributions are in italic). In future work, the user model can be extended with more information, such as portfolio. The identification category contains personal information about users. The preference is layout, color scheme and language preferences. The security contains username and password.

The browsing goal concept represents browsing aims of users and it is divided into two sub-concepts: short-term browsing goal and persistent browsing goal. The short-term browsing goal indicates the current information needs of the user. The persistent browsing goals are long-term goals of the user, which are motivated from long-term interests. Such as, a user interested in politics, probably likes politic related pages. In our ongoing work, browsing goals are automatically defined based on the semantic context found on the page. Users are only required to select appropriate browsing goal from their browsers. The interest category represents browsing interests of users that can be understood from bookmarked pages and accessed semantic hyperlinks. Thus, the interest category is divided into bookmark (interest to a webpage) and browsing interest (interest to a semantic concept). Users can explicitly enter interest values to recognized semantic instances from their browsers. The expertise category represents expertise of users for a semantic instance. Users can explicitly enter expertise values to the semantic instances recognized on the Web page.

Additionally, in order to implicitly understand the activities of users, we introduce the Browsing Behavior concept. The browsing behavior has *browsing_level* and *browsing_type* properties. *Browsing_level* (*very active*, *active*, *passive*, *inactive*) is

the number of clicks made by a user in a browsing session. According to Cove and Walsh [1], and Bawden [2], different browsing strategies exist and we use browsing_type (*directed*, *semi-directed*, *un-directed*) to represent this. In our model, browsing strategy of the user is implicitly understood from their interactions with SemWeb. When the user has a short-term browsing goal, it is assumed that user is looking for a defined piece of information and browsing_type is set to “*directed*”. This probably accounts for perception of search browsing. When the user has a browsing interest or has bookmarked current Web page, then it is assumed that the user is looking for items of interest and browsing_type is set to “*semi-directed*”. This probably accounts for perception of general purpose browsing. When the user does not have short-term browsing goals or browsing interests, browsing_type is set to “*undirected*”. This probably accounts for perception of serendipity browsing.

4 Conclusions and Future Work

In this paper, we have presented our ongoing work for the novel personalized Semantic Web Browser (SemWeb). Our main contribution is the integration of a new behavior-based and an ontology-driven user modeling architecture into the Semantic Web browser. As a result, SemWeb provides semantic and adaptive hyperlinks on different web sites and user profiles can be easily extended with semantic metadata coming from browser. We have also introduced a new user model, which uses the user’s browsing behaviors for adaptation and it can be applied to different domains.

In our ongoing work, adaptive hypermedia is achieved on the recommended semantic links based on the user model and found concept instances on the page. In addition, based on different browsing types and expertise, we are planning to provide adaptive links. For instance, if it is directed browsing, show related links according to short-term browsing goals. If it is semi-directed browsing, use most recently added interests to supply related links. If it is un-directed browsing, make use of semantics. When a link is requested by a novice user, provide links to Wikipedia pages. When the user is an expert, provide detailed semantic links. Also, link sorting and link annotation can be done based on interest ratings, goal priorities, and browsing levels.

References

1. Bawden, D.: Information Systems and the Stimulation of Creativity. *J. of Information Science*, vol. 12, pp. 203-216, (1986)
2. Cove, J., Walsh, B.: Online Text Retrieval via Browsing. *Information Processing and Management*, vol. 24, pp. 31-37, (1988)
3. Carr, L., Hall, W., Bechhofer, S., and Goble, C.: Conceptual Linking: Ontology-based Open Hypermedia. *International World Wide Web Conference*, pp. 334-342, (2001)
4. Dzbor, M., Domingue, J., and Motta, E.: Magpie – towards a semantic web browser. *ISWC*, (2003)
5. Razmerita, L., Angehrn, A., Maedche, A.: Ontology-based User Modeling for Knowledge Management Systems. *International Conference on User Modeling*, pp. 213-217 (2003)
6. Yudelson, M., Gavrilova, T., Brusilovsky, P.: Towards User Modeling Meta-Ontology. *International Conference on User Modeling, LNAI*, vol. 3538, Springer, (2005)
7. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V.: GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. In *Proceedings of 40th Anniversary Meeting of the Association for Computational Linguistics*, (2002)
8. ECS Ontology, <http://id.ecs.soton.ac.uk/docs/>