A Unified Ontology-Based Web Page Model For Improving Accessibility

Ruslan R. Fayzrakhmanov
Max C. Göbel
Wolfgang Holzinger
Bernhard Krüpl
Inst. of Information Systems, DBAI Group
TU Wien, Austria

Robert Baumgartner
Lixto Software GmbH
Vienna, Austria

ABSTRACT

Fast technological advancements and little compliance with accessibility standards by Web page authors pose serious obstacles to the Web experience of the blind user. We propose a unified Web document model that enables us to create a richer browsing experience and improved navigability for blind users. The model provides an integrated view on all aspects of a Web page and is leveraged to create a multi-axial user interface.

Categories and Subject Descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces

General Terms: Design, Human Factors, Experimentation.

Keywords: Web Accessibility.

1. INTRODUCTION

Web accessibility research is strongly focused on the authoring process (WCAG, WAI-ARIA, etc.) of Web documents. Yet only a small percentage of Web sites implement the proposed standards and new technologies require long adoption times for accessibility. While we fully acknowledge the important work in creating and implementing standards for Web accessibility, we see our focus in those websites that, deliberately or not, do not adhere to these standards and are currently hard to read by visually impaired users.

In the ABBA[1] project, we focus on improving the navigability of a Web document for the visually impaired user. Other interesting readings on the topic can be found at [2, 3]. We propose a highly efficient, multi-axial serialization framework to help the user to get a better cognitive understanding of a Web resource and to be able to navigate on that resource more efficiently.

Central to our navigation framework is the transformation of the original HTML format of a Web document into a rich document representation — encapsulating on top of the parsed DOM tree structure a wide set of implicit visual and semantic page properties in a queryable form. This allows to reason more fully over structure, content and interactive functionality of the document and to deduce a wider set of serializations thereof.

In this paper we describe the transformation process as a combination of a number of structural and semantic enrichers and argue that the expressiveness of the resulting model is superior for document serialization to existing approaches. Finally, we present a multi-axial serialization framework that encapsulates the document model into a simple user interface that supports the visually impaired (VI) user in understanding a document and finding required information more effectively.

[1]The ABBA project (Advanced Barrier-free Browser Accessibility) is sponsored by the Austrian Forschungsförderungsgesellschaft FFG under grant 819563.

WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.
ACM 978-1-60558-799-8/10/04.

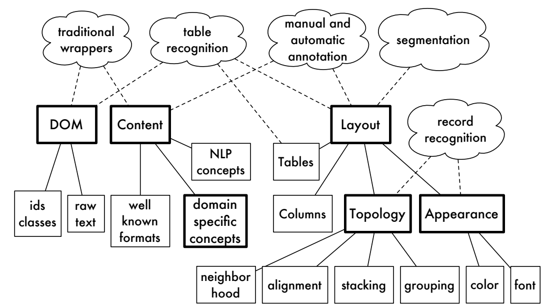

Figure 1: Unified Ontological Model

2. THE UNIFIED ONTOLOGICAL MODEL

Our core innovation towards improving Web navigability for VI users is the introduction of a unified ontological model that i) integrates the descriptions of all aspects of a Web page into one model, and ii) hides away all implementation details from its authoring process. The unified model is a conglomerate of several specialized ontologies providing us with a rich set of concepts for the description of the building blocks on a Web page (words, images, form elements, but also more abstract patterns such as columns) with their properties (coordinates, size, font formatting) and relations (topological, semantical).

Figure 1 shows a top-level outline of the main constituents of the unified model. Boxed elements show sub-ontologies and concepts, cloud shapes indicate algorithms and heuristics that work on them.

For instance, the DOM ontology fragment contains data that is fetched from the live page model of a Web browser. It enables traditional DOM tree based segmentation. The layout ontology fragment describes the visual appearance of page elements like colors and font styles, and their topological relations: relative positions of elements, alignments between elements and several kinds of aggregations of elements for building and referencing more abstract semantic blocks. It also covers more abstract layout elements such as tables, column layouts, lists and other common Web design patterns. In the context of accessibility, the layout fragment is especially interesting, because it contains all those concepts that are completely inaccessible to VI users, but which nevertheless play an important role in organizing the underlying text elements. The content fragment deals with named entities such as place names, proper names, and well known formats as dates or prices. This fragment is also extended by domain-specific concepts and vocabularies.

The cloud shaped algorithms in the diagram are *enrichers* for the unified model. They gather their input data by querying the model and bootstrapping their output back into the model, using the conceptual language of the unified model. For instance, a table recognition enricher can work on *DOM* and *Layout* features in parallel by reading their instance data. Enrichers can also be rule based. The unified model has the advantage to provide a common language for all the different aspects of a Web page, and allows to use a reasoner (e.g. Pellet) to discover further knowledge from the document. Rather than navigating on a DOM tree, each navigation request can be translated into a SPARQL query that is executed on the model of the given page. The benefit of the unified model is flexibility, an integrated view and its extensibility for additional semantic information.

3. NAVIGATING THE MODEL

The unified ontological model enables us to view a document as a rich and well-structured taxonomy of concepts. This taxonomy can also be seen as a graph over the document, where nodes represent the individual building blocks and edges the concept relations between the respective entities. Navigating the document now reduces to finding an optimal path through the document graph. In our work, we refer to such paths as *axis*. Similarily as a path in a graph implies a strict ordering on its nodes, an axis is an ordering function over a set of input entities of the document model.

There are many different orderings available—depending on the usage scenario, the user, the input document, and a number of other factors. Axes can be of any length and may contain any element, from basic text to complex figures. For instance, an axis can link the entries in a navigation menu or the articles on a news site. Simpler axes mimic the behavior of existing screen readers, ordering text elements in the reading order they appear on the page; or allow the user to jump from hyperlink to hyperlink. To capture the concept of using many axes simultaneously, we call our approach *multi-axial*.

The advantage of the multi-axial model is that an axis simply expresses a relation among entities and therefore can be extended arbitrarily. Landmarks—knots in the axis model where several axes intersect—help users to get a better cognitive understanding (mental model) of a Web page and also allow for a better orientation. At any point, the user is free to switch to a different axis. By changing axes on a knot, the user creates a continuous serialization flow, taking her across different conceptual sections of the document. In Figure 2 we depict two of the axes (navigation, article) that are available to the user on a news Web page.

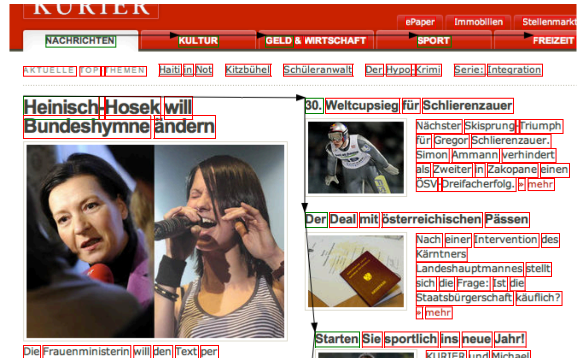

Figure 2: Example of a multi-axial serialization

4. IMPLEMENTATION AND VALIDATION

Text UI prototype. To test the adequacy of the model, we developed a prototype based on the Webkit rendering engine. Internally, a document gets loaded, rendered and then transformed to a RDF document according to our unified model. This document is then augmented with additional semantic knowledge extracted by a number of enricher modules. A first set of enrichers that provide different kinds of segmentations already has been implemented, and we plan to add more enrichers that stem from our previous work in WIE [1], including table understanding and natural language processing modules. We also developed a browser based annotation tool that allows us to simulate the results of future enrichers by tagging a number of test websites with concepts from our model.

For navigation, a set of axis relations is defined on the RDF output. Having finished this initial setup, the prototype chooses an entry point to the document and waits for navigation requests by the user. Navigation possibilities are: i) change axis, ii) go forward, iii) go backward, iv) read, v) orientation. Each request is translated into a query on the model, and the response is then converted to voice output.

Validation. We evaluated our model in respect of i) its capability to model various Web pages, ii) the extensibility when adding new enrichers, and iii) the axes UI it makes possible. We successfully transformed Web pages from four different genres (news, shopping, forum, social site) and used our enrichers and annotation tool to create a number of axes. The evaluation of the axes interface is out of the scope of this paper. Initial tests with probands from the local blind community encourage us to start with the addition of more enrichers. We could already see that our unified model helped us to get a more integrated view of the problem.

5. REFERENCES

[1] Gatterbauer W., Bohunsky P., Herzog M., Krüpl B., Pollak B.: Towards Domain-Independent Information Extraction from Web Tables (2007), In *Proc. of the Sixteenth International World Wide Web Conference, Banff (WWW2007)*

[2] Takagi, H. et al.: Analysis of navigability of Web applications for improving blind usability (2007), In *ACM Transactions on Computer-Human Interaction, 14(3)*

[3] Bigham, J., Prince, C., Ladner, R.: WebAnywhere: a screen reader on-the-go (2008), In *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*