Mining Models for Non-Visual Web Transactions

Jalal Mahmud
Department of Computer Science
Stony Brook University
Stony Brook, NY 11794, USA
jmahmud@cs.sunysb.edu

ABSTRACT

Web transactions (e.g. buying a CD player on the Web) typically involve a number of steps spanning several pages. This task gets strenuous when the Web is accessed non-visually (e.g. when the user is a visually handicapped individual). But usually one needs to browse only a small fragment of a Web page in a transactional step such as a form fill-out, selecting an item from search results, etc. Identifying and presenting such segments from a Web page can overcome the information overload problem when accessing the Web using non-visual interaction modalities like speech. Based on the aforementioned observation I have developed a transactional model that delivers only the "relevant" page fragments at each transactional step, thereby reducing the information overload. In my prior research, such transactional models were learned using a supervised learning approach. Supervised learning requires manually labeled training examples, consequences being that it is nether scalable nor flexible for creating personalized transaction models for arbitrary Web sites. In my PhD research, I am exploring automatic mining of transaction models from unlabeled transactional sequences. My approach is centered on leveraging contextual information of a hyper-link, geometric segmentation of a Web page, and clustering to mine the models. The approach will be scalable but more importantly end-users (with visual disabilities) will be able to create their own "personalized" transactional models for Web sites that they need to use on a regular basis for doing online transactions.

1. INTRODUCTION

The primary mode of interaction with the Web is via graphical browsers, which are designed for visual interaction. As we browse the Web, we have to filter through a lot of irrelevant data. Sighted individuals can quickly segment any Web page and identify the information that is most relevant to them. Individuals with visual disabilities, have to use screen readers [1, 2]. Typically screen readers process a page sequentially but do provide shortcuts to skip over segments and thereby improve browsing efficiency. But nevertheless the information overload problem in some form or the other persists.

This problem is further exacerbated when Web browsing spans several pages as in online transactions, e.g. shopping, registrations, bill payments, etc. In particular, the loss of spatially organized content makes it difficult for blind users to comprehend the sequence of transactional steps. Thus, there is a need for developing techniques that will enable blind users to conduct Web transactions with the same ease as their sighted counterparts.

We capture the two aspects of a Web transaction, namely its operation sequence and content identification by a transactional process model and an ontology respectively. The ontology describes the set of semantic concepts occurring in Web pages, which are considered essential for conducting Web transactions in a particular domain. The transactional process model is a deterministic finite automata (DFA) that captures the set of transactional sequences. Each state, representing an atomic operation in a transaction, is associated with a set of semantic concepts drawn from the ontology. When the model makes a transition to a state during the course of a transaction, a Web page is provided to the state as an input. If the concepts associated with the state are present in the page, then they alone are identified and presented to the user.

In our previous work [6], we demonstrated that coupling content semantics with model-directed navigation facilitated by the process model, can overcome the information overload problem by delivering relevant content at every step of the transaction. We used supervised machine learning technique to learn the transaction model from labeled Web transaction sequences. Sighted users labeled these sequences, as well as instances of semantic concepts in a Web page. We also used a predefined ontology, describing semantic concepts and associated operations.

However, such a supervised approach is not readily scalable to different content domains. Besides the supervised approach makes it almost impossible for blind users to build their own personalized transaction models for Web sites that they have to use regularly like online banks and utilities.

The goal of my dissertation research is to automatically mine transaction models from unlabeled sequences. Unsupervised mining methods address the above limitations of supervised methods.

The approach uses a number of ideas: context surrounding a hyper-link, geometric segmentation of a Web page and unsupervised clustering. The link context (our prior work on contextual analysis appears in [4]) coupled with geometric segmentation [4] is used to retrieve the Web page segment. These segments are clustered based on generic features. The clustered segments represent the semantic concepts in the ontology. A classifier is automatically learnt for each such concept. The process model is learnt using two new ideas – progress and completion equivalence of transactions.

.

2. THE APPROACH

The following subsections briefly describe my approach.

2.1 Label Generation

An unlabeled transaction sequence is a sequence of links followed by the user to perform a Web transaction. Here each link click event is regarded as an atomic operation. First, we have to retrieve the label of the concept associated with each link. Towards that, I use context analysis and geometric segmentation (the algorithms are described in [4]) to get the Web page segments containing semantically related items. Such segments are called **concept segment** (since they contain semantic concepts). Once the concept segments are retrieved, I apply a clustering technique (e.g. greedy clustering) among them to put similar segments into the same cluster. These clusters denoted as $C1, C2, ..., Cn$ represent concepts in the ontology

I then use the concept labels (retrieved after clustering) to label each atomic operation in the transaction sequences.

2.2 Learning Concept Models

I have developed SVM based concept models using features derived from concept segments. Currently words, bigrams, trigrams, stemwords, stem bigrams, stem trigrams, and presentational patterns (e.g. link, text, image) are used as features for learning the statistical concept model.

2.3 Learning Process Models

In my previous work [6], I used DFA learning to learn a process model for online shopping [6]. It uses traditional language equivalence and does not capture the transactional semantics. There are two drawbacks with learning DFAs. The size of the model grows with the number of training examples. Moreover, DFA learning requires sizable number of negative examples which are often difficult to obtain.

I am developing a process model learning algorithm. Key aspects underlying the learning are as follows: The algorithm only uses complete transaction sequences. These denote completed transactions. It also utilizes two new notions – **Progress Equivalence** and **Completion Equivalence**. The former is based on the observation that a user does not make "progress" by doing repeated sequences of transaction operations. For example, the transaction sequences $< SelectItem, AddToCart >$ and $< SelectItem, AddToCart, SelectItem, AddToCart >$ make equivalent progress. Completion Equivalence conveys that final states are equivalent no matter what transactional paths are used to reach them.

3. PROJECT STATUS

Some of the main components of the system have been developed and tested. For example, geometric segmentation and context analyzer has been extensively tested for over two dozens of Web sites and the evaluation results appear in [4]. Experimental performances of our supervised models and the integrated system are reported in [6]. SVM based concept models are trained and tested for few Web sites. Extensive evaluation is a work in progress. Implementation of the unsupervised mining method is underway. The integrated system for conducting Web transactions will be evaluated by blind as well as blind and deaf students at Helen Keller Services for the Blind and Helen Keller National Center in NY.

4. RELATED WORK

This work has broad connections with research in non-visual Web access, Web page segmentation, process model learning, etc. Blind users access the Web using screen-readers, such as JAWS [1] and IBM's Home Page Reader [2]. However, they do not help find the relevant segment containing semantic concept in that page.

Substantial research has been done on segmenting Web documents [7, 3]. Semantic partitioning of Web pages has been described in [5]. These techniques are either domain specific [7] or depend on fixed sets of manually specified rules [3]. In contrast, my geometric segmentation is fully automated and scalable over domains and does not depend on manually specified rules or domain knowledge.

Learning process models from user activity logs is related to research in mining workflow process models (see [8] for a survey). However, our current definition of a process is simpler than traditional notions of workflows.

5. CONCLUSION

I have described my ongoing PhD research on a model-driven approach for doing Web transactions using non-visual interaction modalities. The models will be automatically mined from unlabeled transaction sequences using unsupervised clustering, classification, and my prior research on web page partitioning and context analysis.

The transaction system that will emerge from my PhD research will push the state-of-the-art in assistive web browsing technology. In particular blind users will be able to personalize the system for Web sites that they need to use on a regular basis and conduct transactions on these sites with the same ease as their sighted counterparts.

6. REFERENCES

[1] JAWS Screen Reader. http://www.freedomscientific.com.

[2] C. Asakawa and T. Itoh. User interface of a home page reader. In *ACM Intl. Conf. on Assistive Technologies (ASSETS)*, pages 149–156. ACM Press, 1998.

[3] D. W. Embley, Y. Jiang, and Y.-K. Ng. Record-boundary discovery in web documents. In *ACM Conf. on Management of Data (SIGMOD)*, 1999.

[4] J. Mahmud, Y. Borodin, and I. V. Ramakrishnan. Csurf: A context-directed non-visual web-browser. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, 2007.

[5] S. Mukherjee, G. Yang, and I. Ramakrishnan. Automatic annotation of content-rich html documents: Structural and semantic analysis. In *Intl. Semantic Web Conf. (ISWC)*, 2003.

[6] Z. Sun, J. Mahmud, S. Mukherjee, and I. V. Ramakrishnan. Model-directed web transactions under constrained modalities. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 447–456, 2006.

[7] H. Takagi, C. Asakawa, K. Fukuda, and J. Maeda. Site-wide annotation: Reconstructing existing pages to be accessible. In *ACM Intl. Conf. on Assistive Technologies (ASSETS)*. ACM Press, 2002.

[8] W. van der Aalst and A. Weijters. Process mining: A research agenda. *Computers and Industry*, 53:231–244, 2004.