

Ion Juvina & Herre van Oostendorp

Enhancing internet experience of visually impaired persons by means of dynamic highlighting and selective reading

Keywords: web navigation, visually impaired persons, cognitive modeling, intelligent agents, dynamic highlighting, selective reading

This paper proposes a research-based tool to assist visually impaired persons (VIPs) in using the Internet via screen readers. The proposed tool is inspired by research on modeling web use and model-based highlighting. This tool assists VIPs in selecting goal-relevant information on web pages. A computational cognitive model simulates the VIPs' Internet use. An intelligent agent capable of dynamic highlighting and selective reading based on efficient machine learning algorithms runs alongside the (simulated) user. The agent learns from interacting with the cognitive model and the information space. This agent is implemented in an adaptive interface that takes, expands and updates a user goal, finds goal-relevant information and suggests it to the (simulated) user in an appropriate way. The proposed tool could be applied in situations that require handling information overload with limited perceptive and cognitive capabilities.

Introduction

Visually impaired persons (VIPs) accessing the Internet¹ via *screen readers*² have difficulties locating goal-relevant information (Jones, Farris, Elgin, Anders &

Johnson, 2002). Using the Internet naturally involves revisits to certain pages and one of the most frequent user actions is pressing *the back-button* (Cockburn, Greenberg, Jones, McKenzie & Moyle, 2003). Re-visitation is not only a means to correct for superficially processed or forgotten information, it is also a way to involve information contexts in judging the relevance of a particular piece of information. According to Wen (2003), users employ exploration and re-visitation in order to get acquainted with the context of a particular piece of information, which is eventually useful in judging the value of that particular information. Due to re-visitation, VIPs have to listen redundantly to large amounts of content and options. This slows down the process and adds information load that must be handled with users' limited cognitive resources. Users' cognitive overload increases and the quality of their Internet experience decreases. This might explain why only 21% of VIPs have access to the Internet, whereas 57% of sighted persons have access to the Internet (Gerber & Kirchner, 2001). But, even when basic access to the Internet is ensured, questions arise regarding the quality of users' experiences with the Internet. How much do they really benefit from using the Internet? Are they fully capable of taking advantage of the whole functionality? (Di Blas, Paolini & Speroni, 2004)

Current screen readers read out in speech or Braille characters the content and options on web pages in a

non-selective way. They do not provide enough support in discriminating between various information types. Sighted persons have size, colors, position, shape, motion etc. as aids to discriminate between various types of information. Screen readers have only phonetic cues (e.g. male vs. female voice) and verbal cues (e.g. the word 'link' is added after each link). Moreover, there is no support in discriminating goal-relevant information from 'noise'. Information is presented serially and not in parallel as in visual interfaces. Visual and contextual cues are either absent since they cannot be translated in a textual form or difficult to retrieve because re-visitation is non-selective. Selective reading as an alternative to exhaustive reading would be a natural solution. This requires adding 'intelligence' to screen readers. An intelligent agent could be working together with the screen reader prompting the user with goal-relevant selections.

We have developed a framework for modeling web use and generating model-based navigation support (Juvina & van Oostendorp, 2004, 2006), which builds on existing theories and models of reading comprehension (Van den Broek, Young, Tzeng & Linderholm, 1999) and Web use (Pirolli, 2005; Kitajima, Blackmon & Polson, 2000). This framework together with one previous study is briefly presented in this paper. On basis of this framework we discuss how VIPs can be assisted in using the Internet by means of the latest technology. Specifically, we propose an intelligent agent capable of dynamic highlighting and selective reading. The proposed agent is presumably able to interpret a given user goal, find goal-relevant information and suggest it to the user in an appropriate way while running alongside the user.

Background

Tools have been designed to support people in using the Internet, in particular, in selecting goal-relevant information: *orientation* tools such as maps and indexes allow

zooming in a particular area of the information space and, thus, support rapid selection of relevant information; *navigation* tools such as links and menus allow users to explore sequentially the structure and contents of the information space; and *search* tools try to find the best match between a particular information request of the user and the available information regardless of its location or context. Since these tools are rather complementary than mutually exclusive (Teevan, Alvarado, Ackerman & Karger, 2004), efforts are being made to integrate them.

For example, a tool that tries to integrate navigation and search is proposed by Olston and Chi (2003). Their tool called ScentTrails highlights links on webpages to indicate paths to search results. Links are highlighted according to their relevancy to keywords entered by the user (Olston & Chi, 2003).

Significant work has been recently devoted to building tools for assisting VIPs in their Internet use. One approach is concerned with making auditory interfaces as similar as possible to visual interfaces. Thus, tools and techniques proven useful for sighted users can be converted so as to be useful for VIPs as well (Frauenberger, Stockman, Putz & Holdrich, 2005). For instance, an analog of visual scanning can be performed on audio data by speeding up the reading process (Hurst, Lauer, Burfent & Gotz, 2005). Auditory icons (Petrie, Morley, McNally, O'Neill & Majoe, 1997) and audio feedback indicating the user's location can be provided by a combination of speech and non-speech sounds (Strain & McAllister, 2005). Another approach focuses on treating the content of Web pages (Yu, McAllister, Kuber, Murphy & Strain, 2005) and modifying them so as to make them accessible for VIPs, for instance, by summarization (Zajicek, Powell & Reeves, 1998).

When such tools are based on theory and research, they have a higher chance to be successful and they can be generalized beyond their initial application domain.

In turn, attempts to build and implement such tools can lead to validation or refinement of the theories that inspired them in the first place.

Modeling Web use

Information processing theories of text comprehension and human-computer interaction, such as the *Construction-Integration* theory of text comprehension (Kintsch, 1998) and the *Information Foraging* theory (Pirolli & Card, 1999), have inspired scientists and practitioners to build models of Web use and support tools. In particular, these theories have inspired computational models of reading (Van den Broek, Young, Tzeng & Linderholm, 1999) and Internet use (Kitajima, Blackmon & Polson, 2000; Pirolli & Fu, 2003). These cognitive models were developed not only with the aim of testing and refining the associated theories but they were also meant as tools for practitioners.

CoLiDeS – A Comprehension-based Linked model of Deliberate Search (Kitajima, Blackmon & Polson, 2000) measures relatedness of a particular screen object to the user's goal (information scent) based on three factors: semantic similarity, frequency and literal matching. The semantic similarity measure is based on Latent Semantic Analysis (LSA) (Landauer, Foltz & Laham, 1998). The meaning of a word is determined by what surrounds it (term by document co-occurrences). A semantic space is built based on singular value decomposition (a kind of factor analysis).

For example (Zampa & Lemaire, 2002): if word 1 (e.g. bike) co-occurs with words (e.g. handlebars, pedal, ride) that co-occur with word 2 (e.g. bicycle), and word 1 does not co-occur with words (e.g. flower, sleep) that do not co-occur with word 2, then words 1 and 2 are considered quite similar.

CoLiDeS models mainly the ideal situation of forward linear navigation; backtracking steps are considered erratic actions. However, using the Web naturally involves revisitation of pages (Cockburn, Greenberg, Jones, McKenzie & Moyle, 2003). Howes, Payne and Richardson (2002) argue that a search strategy is required in addition to an ability to follow label semantics, because in real Web navigation, label semantics are rarely sufficient to guarantee that users will navigate directly to the location of a goal without exploration of other parts of the search space.

Our own attempts at modeling Web navigation behavior have been reported elsewhere (Juvina & van Oostendorp, 2004; Juvina & van Oostendorp, 2006) and are only summarized here. In characterizing users Web navigation behavior we have used the distinction syntactic–semantic–pragmatic, adapted from linguistics. Users' navigation takes place in an information space and has a particular structure. In pursuing their goals, users follow *paths* through this space. Paths are linked information elements (e.g. Web pages) that are deliberately selected by the user. One can identify syntactic features of navigation paths and syntactic roles of information elements in a path: order of selections, arrangement of pages, pages having specified roles (home, index, hub) etc. Navigation paths can also be semantically characterized by, for instance, page contents. Finally, relating navigation paths to users tasks and goals is a way to characterize user navigation behavior in a pragmatic way (Juvina & van Oostendorp, 2004).

Navigation metrics indicating aspects of how the navigation graph is structured were labeled syntactic metrics. Metrics indicating content-related aspects of web pages were called semantic metrics. Examples of syntactic metrics are: the rate of revisits, the number of cycles and the average connected distance between two nodes (Juvina & van Oostendorp, 2006). Semantic metrics are '*information scent*' and '*path adequacy*'.

Information scent is calculated as a LSA semantic similarity coefficient between links texts and the user's goal (Kitajima, Blackmon & Polson, 2000). Path adequacy is the semantic similarity between a navigation path and a user's goal (Juvina & van Oostendorp, 2004). A navigation path is a succession of links that have been selected prior to a particular moment in a navigation session, and is, in cognitive terms, a memory representation of the user's past selections. Information scent pertains to the incoming information elements, whereas path adequacy indicates the goal-relevance of past selections. Based on both information scent and path adequacy, the user can decide whether a candidate information element is goal-relevant and consistent with past selections.

We have made a few amendments to CoLiDeS, the cognitive model of Web navigation proposed by Kitajima, Blackmon and Polson (2000), and the altered model has been labeled CoLiDeS+. CoLiDeS+ brings in the concept of *path adequacy* as a complement to the concept of *information scent*. By including the user's navigation path in the model, we aimed to making the model accountable for the role of contextual information in selecting specific navigation actions and further on for user's spatial cognition involved in Web navigation. On basis of the correlation between path adequacy and task performance we were able to show (Juvina, van Oostendorp, Karbor & Pauw, 2005) that path adequacy was a relevant variable describing user navigation behavior.

Model-based navigation support

We have experimented with navigation support in the form of suggestions (highlighting) generated based on simulations of the CoLiDeS+ model (Juvina, van Oostendorp, Karbor & Pauw, 2005). A number of Web tasks were simulated with CoLiDeS+. The basic idea of this study is that a positive effect of the suggestions would provide empirical support of the plausibility of

the model. Performance-oriented tasks with well-defined goals (example: 'what type of food can be included in a healthy diet?') were used. They required navigating through Web-based documents in order to make sense of different but linked pieces of text.

The results of simulations were successful paths, i.e. successions of links leading to the target pages, and 'dead-ends', i.e. pages that are not linked with the target pages, making it necessary for the user to backtrack. Based on these results of the simulations, two types of suggestions were generated:

- link suggestions – when a link contained in a successful path was visible on the screen, the user received the suggestion *Click on* <link label>;
- path switch suggestions – when a 'dead-end' page was downloaded, the user got the suggestion *Go back*.

Participants were randomly assigned to two conditions: a control condition and a support condition. Participants receiving support were informed that suggestions were generated by a robot, they were meant to help with task execution, and they were not mandatory: participants could follow them or not at their own discretion. Suggestions were provided via voice.

Providing such navigation support made a significant difference in users' navigation behaviour and task performance:

- The number of navigation steps was lower in the support condition (with suggestions) than in the control condition ($t_{27} = 3.86$; $p = 0.001$). It took an average of 30 steps to execute a task in the control condition and only 19 steps in the support condition.
- The average duration of tasks was shorter in the support condition than in the control condition ($t_{27} = 2.16$; $p = 0.04$). It took an average of 10.26 minutes to complete a task in the support condition and 12.49 minutes in the control condition.
- Task performance was significantly higher in the

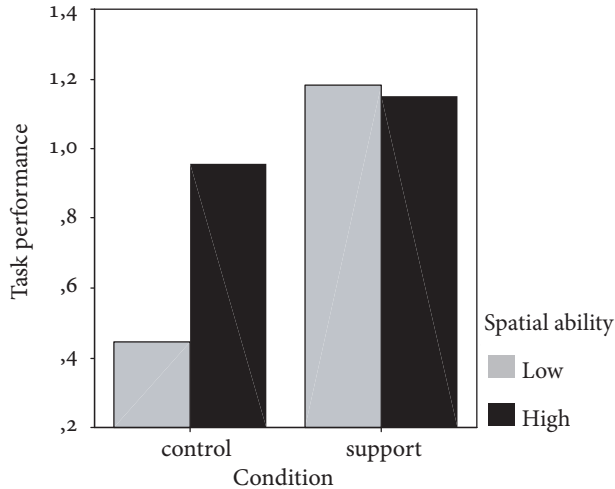


Figure 1. The effect of providing navigation support on task performance for users with low and high spatial ability

support condition (mean = 1.16) than in the control condition (mean = 0.68) ($t_{27} = 2.16$; $p = 0.04$).

- The difference in task performance induced by navigation support was checked separately for low and high spatial ability participants (Figure 1). Participants with low spatial ability had a significantly higher performance in the support condition than in the control condition ($t_{12} = 2.27$; $p = 0.04$). Task performance of participants with high spatial ability was not significantly different between conditions ($t_{12} = 0.73$; $p = 0.48$).

IONS-VIP: a model-based agent for dynamic highlighting and selective reading

As shown in previous research (Juvina & Herder, 2005) and briefly presented above, providing navigation support derived from the model developed had a positive influence on users' perceptions, users' navigation behavior, and task performance. These studies form the direct

impetus for IONS-VIP,³ the tool proposed in this section.

On basis of this framework we want to develop a tool that can assist VIPs in an intelligent way. The attempt to add 'intelligence' to screen readers generates some interesting research and design problems. Basically, the screen reader should be complemented with a software agent that is capable of dynamic highlighting and selective reading, i.e. the agent should be able to predict what a particular user would need to select and read at any particular moment during task execution.

Cognitive modeling component

First, a basic understanding of how VIPs use the Internet is necessary. Some aspects are common for VIPs and sighted persons, whereas others are specific. For example, all users represent their goals and the information space and activate contextual information and background knowledge when judging the relevance of incoming information. However, while sighted users are able to make use of visual cues in selecting the relevant information, VIPs have to rely on sequential and non-selective input of screen data.

Visual and spatial aspects are essential for sighted persons' Web use. Empirical results show a positive correlation between *spatial ability* and Web task performance (Chen & Rada, 1996; Juvina & van Oostendorp, 2006). Spatial terms are extensively used in users verbalization during Internet use (van Hooijdonk, Maes & Ummelen, this issue) and users report maintaining a *sense of location* during their Web navigation (Teevan, Alvarado, Ackerman & Karger, 2004). There is also evidence from cognitive neuroscience research (Gazzaniga, Ivry & Mangun, 1998) showing that humans represent information in terms of *what* (content) and *where* (location). For sighted users, information of type *where* could literally mean visual location. For VIPs *where* information could refer to: temporal position of a

particular information element in a sequence, category membership, etc. Thus, both semantic and spatial (syntactic) processes are involved in using the web. Moreover, semantic and syntactic processes are influencing each other. For example, a higher semantic relevance (information scent) might be assigned to an information element just because it is perceptually more salient than other elements (Tamborello & Byrne, 2005). In the case of VIPs using the web via screen readers, perceptual salience refers to physical properties of the reading voice or position of a particular item in a reading sequence.

All these aspects referring to perceptual and cognitive processes involved in using the web can be incorporated in a cognitive model. Developing a model of VIP's use of the Internet allows understanding basic mechanisms such as goal representation and specifying several key variables such as working memory capacity. A computational cognitive model runs as a computer program and simulates users' task execution. Its output can be used as input for an intelligent agent that provides support with task execution. See the previous section for an empirical demonstration.

Using computational cognitive models as substitutes for users (Ritter, Baxter, Jones & Young, 2000) and as base for building Internet tools is not the only option and definitely not the easiest one. It is beneficial though in a number of cases:

When tools are expected to possess some sort of autonomy and to interact with one another or with users in real time (intelligent agents), they need a process model of use in order to interpret user actions and react accordingly.

When tools are required to preserve users' experiences and to intervene in an unobtrusive way, they need to *know*, *learn* or *predict* certain aspects of users' behavior.

When tasks to be supported are complex, weakly defined, and knowledge-intensive, the supporting tool needs to perform a series of *informed guesses* about what the user is trying to do and what type of support is needed.

Machine learning component

Machine learning algorithms can be used to find the most efficient and effective solutions to the task in real time. These algorithms can be combined and form an intelligent agent capable of dynamic highlighting and selective reading. The agent takes input from the cognitive model (or the user) and explores the information space in searching for information that is most likely to be relevant to the user's goal, given the user's progress with task execution. Specific algorithms are proposed here for (1) computing semantic similarity, (2) computing optimal information paths for emphasizing hyperlinks (reinforcement learning spider), (3) summarizing text fragments with goal-relevant keywords, and (4) user tracing.

Computing semantic similarity. In order to identify relevant information on web pages, machine-learning algorithms (van Kuilenburg, Wiering & den Uyl, 2005; Cilibrasi & Vitanyi, 2005) can be used for computing semantic similarities between user goals and texts on web pages (e.g. link labels).

Reinforcement learning (RL) spider (computing optimal information paths). Based on a neural network that computes semantic similarities between the user's goal and texts on web pages, a RL algorithm (Kaelbling, Littman & Moore, 1996) can learn paths through the information space that maximize the semantic similarity with the user's goal (Joachims, Freitag & Mitchell, 1997).

Summarizing text fragments with keywords. Although sighted users can easily skip text fragments if they are not interested in them, a VIP usually has to wait for the screen reader to finish reading each text fragment, costing more human effort than necessary. A solution to this problem is to make a summary of each text fragment (Mani & Maybury, 1999) and to give the summary using few keywords to the user who can then choose whether to hear the whole fragment or not. For summarizing text fragments with few keywords a search algorithm can be

used for finding keywords that are highly similar to both the user goal and the original text fragment.

User tracer. All the techniques described above depend on the user's goal being known to the system in sufficient detail. The simplest way to acquire the user's goal is to have it entered by the user her/himself. However, user-entered goals are usually weakly specified. They can be enlarged by using Wordnet⁴ (Fellbaum, 1998) to generate synonyms and hypernyms of the given words. Another way to expand the user's goal is based on the user's navigation history. For this, the tracer can search for a set of keywords that has the highest similarity score to a number of visited Web pages. The resulting keywords can be added to the goal description provided by the user.

Besides the user's goal a supporting agent needs to know about the user's actual behavior. In terms of web mining, the agent needs to follow the user and learn by observation and imitation. Machine learning can be used to predict the clicking behavior of a user based on semantic and syntactic features of the user's navigation path (Jovina & van Oostendorp, 2004). Since learning from human users is a very slow process requiring many users and many interactions with the system, the cognitive model can first be used as substitute of a human user. In this way, it becomes much easier and faster to train the predictors than with real users. Once the predictor was able to mimic the cognitive model, it can interact with and learn from human users.

User-interface component

Assuming the cognitive model is valid and the machine learning algorithms are robust and efficient, an interface agent capable of dynamic highlighting and selective reading should be designed and tested to see whether support is provided in an effective and satisfactory way. Once the user goal is known, the interface agent should

find appropriate ways to deliver support. For example, one type of support is suggesting that some contents are more goal-relevant than others by highlighting (emphasizing) the relevant contents. Emphasizing can be implemented by varying the properties of the reading voice or by increasing the presentation priorities of the relevant items.

For example, if the user is looking for 'electronics' in the following menu 'books/music/electronics' it will take at least 3 times longer to find it than if the user were looking for 'books'. So 'books' is 'highlighted' by default because of the sequential character of reading. If 'electronics' is to be highlighted, it should be placed on the first position in the menu.

Specific items might become relevant or irrelevant during task execution depending on what the reader has already done. Therefore, triggering reader's attention toward goal-relevant items at the user interface level should be synchronized with reader's progress through task execution. The type of support we propose can be considered in the category of adaptive guidance systems (Brusilovsky, 2001), in which a cognitive model is used as a user model. Delivering adaptive navigation support must be made with full consideration of usability and accessibility guidelines.

For example, changing the order of reading for some information items in order to emphasize goal-relevance should not interfere with the usability principle of *consistency*. In general, information should be kept as much as possible in its original context, the one decided by content providers.

Integrating all three components can make this approach successful. The interface agent should synchronize the cognitive model and the machine-learning agent in order to provide support in an effective, efficient and satisfactory way. A diagram of our proposed tool is presented in Figure 2.

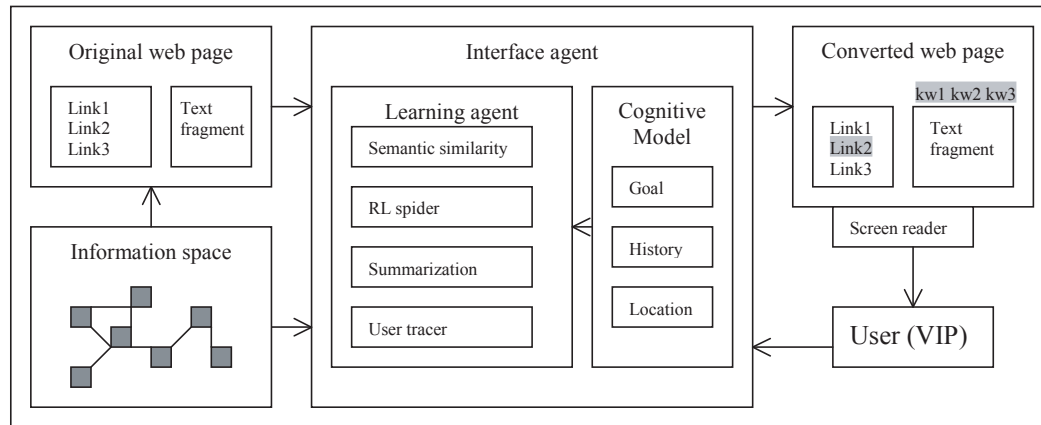


Figure 2. IONS-VIP diagram. Requested Web pages are converted by an interface agent to include navigation support. The interface agent controls the interaction between the cognitive model and the machine learning agent. The cognitive model simulates the user's behavior and gives input to the learning agent. The learning agent takes input from the current webpage, the (simulated) user and the information space and outputs link suggestions and summarizing keywords (Legend: RL – reinforcement learning; kw – key word).

The three components depend on one another: the cognitive model simulates key aspects of task execution and specifies where and when the user could use support; the machine learning agent needs to interact with a (simulated) user in order to collect and learn from data about the user's behavior (user's goal and navigation path); and an interface agent is necessary to ensure a smooth interaction between the learning agent and the (simulated) user. Since both the cognitive model and the learning agent are implemented as computer programs they can be run as many times as necessary until all the relevant aspects of their interaction are fully specified and tested. Certainly, the interface agent ultimately needs to be tested with real users to be proven useful and usable in real-life settings.

Conclusion and discussion

We have presented a process model of Web navigation that it is able to simulate task execution and generate navigation support. The model simulates users' assessment of goal-relevance of text fragments on Web pages and explains the role of contextual information involved in relevance assessments. Providing users with model-generated navigation support had a positive influence on users' perceptions, users' navigation behavior and task performance, especially for users with low spatial abilities.

A model-based tool has been proposed capable of supporting VIPs' Internet use by means of dynamic highlighting and selective reading. The proposed tool has been described in terms of its functionality rather than in technical details. While there is enough theoretic-

cal and empirical ground to make our proposal plausible, its actual implementation is less clear-cut. There are important challenges to be addressed. For cognitive modeling, a task domain as weakly structured and knowledge-intensive as the one proposed here will be a great challenge. A cognitive model of VIPs' use of the Internet needs to handle natural language, large knowledge networks, and a great deal of sub-symbolic computations. These aspects are not part of the traditional work in cognitive modeling research (Gluck & Pew, 2005) but are becoming increasingly prominent in the cognitive modeling community (Pirolli, 2005). A foreseeable problem will be to handle the computational complexity required by on-line running of the algorithms. In general, a scalability problem can be expected for both the cognitive model and the machine-learning agent. A number of development cycles can be conceived, starting with small-scale models running off-line and ending with a fully functional tool able to be plugged in and run alongside standard browsers and screen readers.

The solution proposed in this paper differs essentially from static solutions to improve Web accessibility. Such a static solution can be specifying in advance which information should be emphasized or omitted.

For example, a content author can indicate (by *alt=""*) that an image is non-essential and can be ignored by the screen reader (Easton, 2005).

Such a solution could be conceived for text as well, but its applicability and usefulness are rather limited. Commonly, the same content is used for different tasks and purposes and it is impossible to specify in advance which information elements are (non)relevant. Moreover, information becomes more or less relevant depending on the user's progress in task execution. In other words, relevance is relative to users goals and their progress through their tasks. Therefore, adaptive solutions are rather a necessity than a luxury.

Nevertheless, some intermediary versions of the proposed tool (and the final version as well) can be used in an offline mode in designing and evaluation of user-interfaces for Web applications.

For example, running the agent through a specified Web application for a particular task can detect and repair usability problems such as improper or confusing link labels (Blackmon, Polson, Kitajima & Lewis, 2002), weak coherence among linked pieces of texts (Salm-eron, Kintsch & Canas, this issue), sub-optimal site structure (Miller & Remington, 2002), etc.

The proposed tool can be used for a larger user population than the target population discussed in this paper (VIPs). Partially sighted persons, elderly, and persons engaged in multitasking (e.g. driving while checking route instructions or weather forecasts on the Internet via a voice operated browser) (Velleman & Snetselaar, 2000) are expected to benefit from navigation support of the type proposed here.

Acknowledgement

Marco Wiering (Institute of Information and Computing Sciences) contributed with valuable ideas, particularly related to the connections between machine learning and cognitive modeling and adaptive interfaces.

Notes

1. The terms 'Internet', 'Web' and 'World Wide Web' are treated as synonyms in this paper.
2. A screen reader is a software program which converts screen information to Braille characters and/or speech.
3. The acronym IONS-VIP stands for Internet Orientation, Navigation and Search for Visually Impaired Persons.
4. WordNet® is an online lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory (<http://wordnet.princeton.edu>).

References

- Blackmon, M.H., Polson, P.G., Kitajima, M., & Lewis, C. (2002). Cognitive Walkthrough for the Web. *CHI 2002 Conference*, Minneapolis, Minnesota, USA., ACM.
- Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, 11, 87–110.
- Chen, C., & Rada, R. (1996). Interacting with hypertext: A meta-analysis of experimental studies. *Human-Computer Interaction*, 11(2), 125–156.
- Cilibrasi, R., & Vitanyi, P. (2005). *Automatic Meaning Discovery Using Google*. Technical report Centrum voor Wiskunde en Informatica (CWI), Amsterdam.
- Cockburn, A., Greenberg, S., Jones, S., McKenzie, B., & Moyle, M. (2003). Improving Web Page Revisitation: Analysis, Design and Evaluation. *IT&Society*, 1(3), 159–183.
- Di Blas, N., Paolini, P., & Speroni, M. (2004). 'Usable accessibility' to the Web for Blind Users. *Proceedings User-Interfaces for All (UI4ALL)*, Vienna.
- Easton, B. (2005). *Speaking ALT Text, Access Matters*. Accessed September 12th 2005 from HYPERLINK "http://www.access-matters.com/2005/05/15/speaking-alt-text/".
- Fellbaum, C., (Ed.) (1998). *WordNet. An Electronic Lexical Database*. Cambridge: The MIT Press.
- Frauenberger, C., Stockman, T., Putz, V., & Holdrich, R. (2005). Design Patterns for Auditory Displays. *Human Computer Interaction*, Edinburgh, UK: Springer.
- Gazzaniga, M.S., Ivry, R.B., & Mangun, G.R. (1998). *Cognitive Neuroscience. The biology of the mind*. New York: W.W. Norton & Company.
- Gerber, E., & Kirchner, C. (2001). Who's surfing? Internet access and computer use by visually impaired youths and adults. *Journal of Visual Impairment & Blindness*, 95, 176–181.
- Gluck, K. A., & Pew, R.W. (2005). *Modeling Human Behavior with Integrated Cognitive Architectures*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Howes, A., Payne, S.J., & Richardson, J. (2002). An instance-based model of the effect of previous choices on the control of interactive search. *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*.
- Hurst, W., Lauer, T., Burfent, C., & Gotz, G. (2005). Forward and Backward Speech Skimming with the Elastic Audio Slider. *Human Computer Interaction*, Edinburgh, UK: Springer.
- Joachims, T., Freitag, D., & Mitchell, T. (1997). WebWatcher: A Tour Guide for the World Wide Web. *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*. Morgan Kaufmann.
- Jones, K.S., Farris, J.S., Elgin, P.D., Anders, B.A., & Johnson, B.R.. (2002). Examining Information Searching on the World Wide Web with a Screen Reader: a Verbal Protocol Analysis. *Human Factors and Ergonomics Society 46th Annual Meeting*.
- Juvina, I., & Herder, E. (2005). The Impact of Link Suggestions on User Navigation and User Perception. *UM2005 User Modeling: Proceedings of the Tenth International Conference*. Edinburgh: Springer-Verlag.
- Juvina, I., & van Oostendorp, H. (2004). Extracting semantic and pragmatic information from Web navigation data for user modeling. *EARLI SIG2 Meeting on Comprehension of Text and Graphics*. Valencia, Spain.
- Juvina, I., & van Oostendorp, H. (2006). Individual differences and behavioral metrics involved in modeling Web navigation. *Universal Access in Information Society*, 1–12.
- Juvina, I., Van Oostendorp, H., Karbor, P., & Pauw, B. (2005). *Toward Modeling Contextual Information in Web Navigation*. Cognitive Science Conference, Stresa, Italy. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kaelbling, L.P., Littman, M.L., & Moore, A.W. (1996). Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research*, 4, 237–285.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Kitajima, M., Blackmon, M.H., & Polson, P.G. (2000). A Comprehension-based Model of Web Navigation and Its Application to Web Usability Analysis. *People and Computers XIV* (pp 357–373). Springer.
- Landauer, T. K., Foltz, P.W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259–284.
- Mani, I., & Maybury, M.T. (1999). *Advances in Automatic Text Summarization*. Cambridge: MIT Press.
- Miller, C. S., & Remington, R.W. (2002). Effects of structure and label ambiguity on information navigation. *Human Factors in Computing Systems Conference*. New York: ACM Press.
- Olston, C., & Chi, E.H. (2003). ScentTrails: Integrating Browsing and Searching on the Web. *ACM Transactions on Computer-Human Interaction*, 10(3), 177–197.

- Petrie, H., Morley, S., McNally, P., O'Neill, A-M., & Majoe, D. (1997). Initial Design and Evaluation of an Interface to Hypermedia Systems for Blind Users. *Hypertext '97*. Southampton, UK.
- Pirolli, P. (2005). Rational Analyses of Information Foraging on the Web. *Cognitive Science*, 29(3), 343-373.
- Pirolli, P., & Card, S.K. (1999). Information foraging. *Psychological Review*, 106(4), 643-675.
- Pirolli, P., & Fu, W.T. (2003). SNIF-ACT: A Model of Information Foraging on the World Wide Web. *Proceedings of the Ninth International Conference on User Modeling*. Johnstown, Pennsylvania, LNAI 2702 (pp. 45-54). Berlin, Heidelberg: Springer Verlag.
- Ritter, F. E., Baxter, G.D., Jones, G., & Young, R.M. (2000). Supporting Cognitive Models as Users. *ACM Transactions on Computer-Human Interaction*, 7(2), 141-173.
- Strain, P., & McAllister, G. (2005). Enhancing Web Accessibility for the Visually Impaired Using RSS. *The 19th British HCI Group Annual Conference*. Edinburgh, The British Computer Society.
- Tamborello, F. P. I., & Byrne, M.D. (2005). Information Search: The Intersection of Visual and Semantic Space. *CHI 2005 Conference*. Portland, Oregon, USA: ACM.
- Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search. *CHI 2004 Conference*. Vienna, Austria: ACM.
- Van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The Landscape Model of Reading: Inferences and the On-line Construction of a Memory Representation. In H. van Oostendorp & S.R. Goldman (Eds.), *The Construction of Mental Representations during Reading*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Van Kuilenburg, H., Wiering, M., & den Uyl, M. (2005). A Model Based Method for Automatic Facial Expression Recognition. *Proceedings of the 16th European Conference on Machine Learning (ECML'05)*.
- Velleman, E., & Snetselaar, H. (2000). Site Seeing. *The development of an accessible web site or web based multimedia product for the sighted, partially sighted and the blind*. Bartimeus Instituut, Zeist, The Netherlands.
- Wen, J. (2003). Post-Valued Recall Web Pages: User Disorientation Hits the Big Time. *IT&Society*, 1(3), 184-194.
- Yu, W., McAllister, G., Kuber, R., Murphy, E., & Strain, P. (2005). Improving Web Accessibility Using Content Aware Plug-ins. *CHI'05 Conference*. Portland, OE, USA.
- Zajicek, M., Powell, C., & Reeves, C. (1998). A Web Navigation Tool for the Blind. *ACM/SIGGRAPH on Assistive Technologies*, Los Angeles, USA.
- Zampa, V., & Lemaire, B. (2002). Latent Semantic Analysis for User Modeling. *Journal of Intelligent Information Systems*, 18, 15-30.

ABOUT THE AUTHORS

Ion Jovina is a PhD student at Center for Content and Knowledge Engineering within the Institute of Information and Computing Sciences, Utrecht University, since 2002. His PhD project is titled 'Development of a cognitive model for navigating on the Web'. He has acquired working knowledge in the field of cognitive modeling at Carnegie Mellon University and has received The Best Student Paper Award at the User Modeling conference Edinburgh 2005.
Contact: ion@cs.uu.nl

Herre van Oostendorp is Associate Professor at Center for Content and Knowledge Engineering within the Institute of Information and Computing Sciences, Utrecht University, since 2000. He is a specialist on the areas of human-computer interaction, usability engineering, hypertext comprehension, discourse comprehension and cognitive modeling.
Contact: herre@cs.uu.nl

Contact

Center for Content and Knowledge Engineering
Institute of Information and Computing Sciences
Utrecht University
Padualaan 14
3584 CH Utrecht
The Netherlands