

IONS-VIP: a cognitive model for navigating the Web via screen readers

Ion Juvina

Department of Psychology, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh PA 15213, USA

ijuvina@cmu.edu

ABSTRACT

This paper presents a cognitive perspective on the problem of using the Internet via screen readers. An empirical study shows similarities and dissimilarities between using the Internet via screen readers and via the standard graphical screen. In particular, spatial cognition is an important determinant of web task performance even in the absence of any visual cues. Backtracking behavior has a higher cost and impacts performance in a more negative way when using the Internet via screen readers as compared to the standard use. Selective reading can be used as an addition to the existing functionality of screen readers but more research is needed to prove the utility of such support mechanism. Computational cognitive modeling can be used in conceiving, testing and implementing cognitively informed support functionality in screen readers.

Keywords

Web navigation, Web accessibility, visually impaired persons, cognitive modeling.

ACM Classification

H5.2. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCERE

A *screen reader* is a software program that converts screen information to Braille characters and/or speech. Visually impaired persons (VIPs) accessing the Internet via screen readers have difficulties locating goal-relevant information [12]. Using the Internet naturally involves re-visits to certain pages and one of the most frequent user actions is pressing *the back-button* [5]. 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. Due to re-visitation, VIPs have to redundantly listen 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 the VIPs have access to the Internet, whereas a higher proportion (57%) of the sighted persons have access to the Internet [8]. 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 [6]?

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.

Significant work has recently been 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 [7]. For instance, an analog of visual scanning can be performed on audio data by speeding up the reading process [11]. Auditory icons [21] and audio feedback indicating the user's location can be provided by a combination of speech and non-speech sounds [25]. Another approach focuses on treating the content of Web pages [28] and modifying them so as to make them accessible for VIPs, for instance, by summarization [29].

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 [20]. 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.

This paper argues for approaching the problem of using the Internet via screen readers in a cognitively informed manner. The following section presents an empirical study aimed at understanding specific behavioral aspects involved in using the web via screen readers. The third section presents a cognitive model of using the Internet via screen readers. The last section concludes the paper and proposes some discussion points and opportunities for further research.

EMPIRICAL STUDY

This study is one of a series of studies investigating cognitive and behavioral aspects of Web navigation. Previous studies in this series have shown that web task

performance depends essentially on users' spatial and semantic abilities [15]. The current study is a replication of previous studies in a new context of use, that is, using the Internet via screen readers. It has been hypothesized that:

- Using the Web via screen readers depends essentially on the same cognitive mechanisms as in the case of using a graphical screen. In particular, the correlation between spatial ability and web task performance, found in our previous studies, will be replicated.
- Navigation support generated by a computational cognitive model can be effectively delivered via screen readers by increasing the reading priority of the selected information.
- Supported users will perform better and be more satisfied than controls.

Participants

A number of 29 undergraduate and graduate students at Utrecht University participated in this study. Participants were not visually impaired and had no prior experience in using a screen reader. Participants received financial compensation for their participation.

Design

An equal number of participants (13) has been assigned at random to two conditions: a *control* condition in which participants had to execute two web navigation tasks via a standard screen reader, and a *support* condition in which participants had to execute the same tasks while receiving navigation support generated by a cognitive model in the form of link suggestions delivered via the same screen reader. Three subjects used for the pilot tests have been added to the control condition, so in the end there were 16 subjects in the control condition and 13 in the support condition.

A mental rotation task [24] has been administered to all subjects to measure their spatial abilities. During the navigation tasks a number of navigation actions were recorded for all users, except the three pilots added to the control condition. Pre- and post-task questionnaires were also administered to all subjects.

Materials

A modified version of *www.howstuffworks.com* has been used. This website provides explanations of various phenomena (e.g., how toilets work, how electronic pet fences work). A selection of approximately 200 topics spread on approximately 1000 pages has been used. All pages have been modified as to contain only text-based information about the selected topics; ads and graphics have been removed.

Two realistic tasks were developed based on specifications regarding ecological validity [19,16] and our own previous experience. One task was about sleeping and dreaming and the other one about dieting. The following is the description of the dieting task, as the participants received it:

"With the summer coming in fast, you want to look fit for the upcoming beach season. There are a few factors that would make you loose weight. You could cut down on

eating or take some diet pills. But you are aware of the risks associated with these options. You have tried dieting before but after a while the weight came back. The best option is to loose weight in a sustainable and healthy way.

In short, what you have to do is:

- Find out what the side effects of diet pills are.
- Find out how you can diet in a sustainable way.
- Find out what you should reduce in order to lower the number of calories you take in."

The screen reader "Home Page Reader" developed by IBM has been used in executing the tasks. This is one of the most common tools of this kind on the market.

A manipulation of the website was performed only for the "support" condition and it was hypothesized that this condition would differ from the control condition with regard to user behavior and task outcomes.

A cognitive model of web navigation developed in previous studies [13] was run in parallel with the participants. Based on input from the website and the user, the model generated predictions of user selections. These predictions were offered as link suggestions to the participants in the "support" condition.

As to the way of emphasizing the suggested item, there were a few options. One option was to add a word or a small expression to the suggested item. For example, a link to be suggested would be followed by the words "click here". This option has been proven to be effective in one of our previous studies but not so well received subjectively by users [13]. In addition, in the case of VIPs' Web use, this option adds information load, since all the information comes via the auditory modality. Instead, we have decided to suggest an item by changing its reading priority. Since the screen reader always reads from upper-left to lower-right parts of the screen, increasing the reading priority of an item means moving it in the reading sequence upper and to the left. 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. The assumption behind this was that users would hear the relevant item sooner and would make their selection before the whole screen is read, avoiding information overload and, in cognitive terms, having to deal with simpler representations of the information space. To maintain a minimal consistency of the website's structure, the main menu on each page (i.e., the highest level in the website's hierarchy) was left unchanged.

Procedure

Sessions took at most two hours and consisted of taking the spatial ability test, taking a prior-knowledge test, performing the two navigation tasks, and filling out the post-task questionnaires. For the execution of the navigation tasks, vision impairment was simulated by turning off the computer screen. Instruction and training on how to use the screen reader were provided before the

two navigation tasks were administrated. The order of the two tasks was counterbalanced.

Since the experimenter had not controlled the quality of suggestions in advance, a specific measure was necessary to verify that emphasizing the goal-relevant items was successful, that is, the reading priority of the relevant items was indeed higher in the support condition than in the control condition. The position of goal relevant options in the reading sequence was recorded. For example, if a goal relevant item were placed second on a webpage it would get position 2.

Task performance was measured in two ways:

- directly, by counting the number of target pages participants have arrived at;
- indirectly, based on correctness of answers to a set of multiple-choice questions.

Prior knowledge on the topics of the tasks was also measured by administrating the task performance questionnaire before task execution.

The Spatial Ability test, the Web Navigation Metrics, Perceived Usability, Satisfaction, and Perceived Disorientation were measured with specific tests, questionnaires, and instruments as in our previous studies [14]. A few questions were added to measure the perceived usability of the screen reader. For example: "It was easy to learn how to use this screen reader" and "The screen reader had all the functions and capabilities I expected it to have".

Results

A first set of results suggests that using the Web via screen readers involves the same underlying cognitive processes as using the Web via a graphical screen:

- Spatial ability is correlated with task performance ($r = 0.564$, $p = 0.023$, $n = 16$, control condition).
- Users with high spatial ability employ less re-visitation ($F_{1,23} = 3.02$, $p = 0.095$, $n = 26$, marginally significant).
- The number of re-visits is negatively correlated with perceived usability ($r = -0.545$, $p = 0.004$, $n = 26$) and perceived usability of the screen reader ($r = -0.534$, $p = 0.005$, $n = 26$).
- Task performance is correlated with perceived usability ($r = 0.421$, $p = 0.023$, $n = 29$) and perceived usability of the screen reader ($r = 0.393$, $p = 0.035$, $n = 29$).

There are also results pointing at particularities of using the web via a screen reader:

- There is a high cost associated with visiting many pages via a screen reader: the longer the path (number of pages visited), the higher the chance of perceived disorientation ($r = 0.408$, $p = 0.039$, $n = 26$), the lower perceived usability ($r = 0.657$, $p = 0.000$, $n = 26$), and the lower perceived usability of the screen reader ($r = 0.637$, $p = 0.000$, $n = 26$).

With regard to the effectiveness of our support mechanism, as a result of our manipulation, the position of

an item has been decreased in the reading sequence from an average of 4.88 in the control condition to an average 3.84 in the support condition. That is, in the support condition links have successfully been pushed up. In 9 out of 25 cases the relevant item has been pushed up to the first position in the reading sequence.

Regarding the impact of suggestions on users, an analysis of covariance with task performance (correctness) as a dependent variable, condition as a factor and prior knowledge as a covariate resulted in no significant differences between the two conditions ($F_{1,26} = 0.446$, $p = 0.510$, $n = 29$). There were also no significant differences with regard to the other measure of task performance (number of target pages reached), satisfaction, disorientation, and navigation metrics.

This lack of impact of our manipulation on users can be explained by one or both of these factors:

The magnitude of our manipulation (moving the relevant item with approximately one position up in the reading sequence) was not big enough to produce a significant effect.

Changing the order of items on Web pages produced a break of the initial coherence as established by the authors. This "side effect" of our manipulation presumably interacted with (and cancelled out) the expected positive effect.

While not excluding the first explanation, we tend to favor the second one, based on two related findings:

The correlation between spatial ability and task performance behaved in the same way as in one of our previous studies [13]: the correlation is significant in the control condition and non-significant in the support condition. This means that our manipulation had indeed an impact on users.

Users in the support condition visited a larger set of pages than users in the control condition, while re-visitation was constant across conditions. This could probably be caused by users' efforts to repair the breaks of coherence caused by changes in the initial order of items on Web pages.

The following is an example of manipulation (from our training task set) that might have caused coherence breaks. The original page has the following options:

- What is caffeine?
- Caffeine in the diet
- Caffeine and adenosine
- Caffeine and dopamine

The manipulated page has an increased reading priority of the option "Caffeine and adenosine", causing the items to be presented in the following order:

- Caffeine and adenosine
- What is caffeine?
- Caffeine in the diet
- Caffeine and dopamine

Perhaps it is confusing for the reader to first present very specialized information on the effects of caffeine on the

brain (“caffeine and adenosine”) and only after that to present general information about what caffeine is.

Conclusion and discussion

It can be concluded that the same cognitive mechanisms are employed as in the case of sighted users – representing the information space, operating on it, and making selections based on judgments of goal-relevance (information scent). However, using the Internet via a screen reader is more taxing from a cognitive point of view, and users are more vulnerable to dissatisfaction and disorientation if they have to visit a large number of pages to reach their goal.

Our way to emphasize the goal-relevant items by increasing their reading priority did not have the expected impact on users’ performance and satisfaction. Suggestions were delivered in a way that did not increase the information load of the interface and was unobtrusive, but it might be that other ways of emphasizing could have been more effective, for example, changing the audio properties of the reading voice.

Visual impairment was simulated by turning off the computer screen and asking sighted users to perform Web tasks via a screen reader. This might raise the question to what extent the results of this study can be generalized to real VIPs. The decision to use such a study design was taken based on (besides practical limitations) discussions with experts in the area of Web accessibility. According to these experts [26] most of the VIPs are faced with the same situation: they need to learn using a screen reader because they (partially) lost their vision. However, our sample has a lower age than the VIPs population. It is to be expected that real VIPs have even higher difficulties in using the Web than our participants, due to their increased age. It would be useful to replicate this study using VIPs as participants.

IONS-VIP¹: AN ACT-R MODEL OF NAVIGATING THE WEB

There are theoretical, methodological and practical reasons for developing not only conceptual but also computational cognitive models. According to [9], computational models promote the development and evolution of conceptual theories showing where the theories agree with behavioral data and where they do not; they can be used to understand and test alternative explanatory constructs; and they promote communication

among researchers within and across research areas. In addition, particularly in the field of human-computer interaction, cognitive models can be used as substitutes for users [23] and as basis for building user models in adaptive Web applications. There is already a tradition of attempts to build cognitive models of using the Internet, and the field seems to increase in popularity [22,16,18,4,3].

The research reported in the previous section was inspired by the aim to develop support tools for Web interfaces. In this case, converting our conceptual model in a computational model would be highly opportune, as a computational model can be embedded in a real web application. This section provides a demonstration of such computational model.

For this demonstration, the ACT-R (Adaptive Control of Thought - Rational) cognitive architecture [1] and its associated modeling environment - ACT-R6 [2] have been used. ACT-R has validated means to model the key concepts of Web navigation behavior. Thus, ACT-R aural module can handle sequential information input as in the case of using the Web by VIPs. Information is coded based on its content (“what”) and location (“where”). The declarative memory module is appropriate to model the dynamic representation of the information space being navigated, since information elements (chunks) can be stored, modified and retrieved according to cognitively plausible sub-symbolic mechanisms such as spreading activation, decay, frequency and recency. Symbolic rules from procedural memory fire depending on the state of the contextual information stored in buffers. These rules issue appropriate actions that have internal and/or external consequences.

In order to achieve its demonstrative purpose, the model presented here interacts with a small scale website and uses an external tool – Generalized Latent Semantic Analysis (GLSA) – to represent background knowledge and calculate semantic similarities. The website is a simplified version of How Stuff Works (<http://www.howstuffworks.com/>). As its title suggests, this website provides explanations of various phenomena such as hurricanes, hypnosis, etc. In this modified version, information on Web pages has been reduced to link labels and each label is composed of a single word. Starting with the home-page, Web pages are presented to the model one after another depending on the model’s selections. For instance, Figure 1 shows the homepage as it is presented to the model. The model selects one of the presented link labels causing the corresponding next webpage to be displayed.

¹ IONS-VIP stands for Internet Orientation, Navigation and Search for Visually Impaired Persons

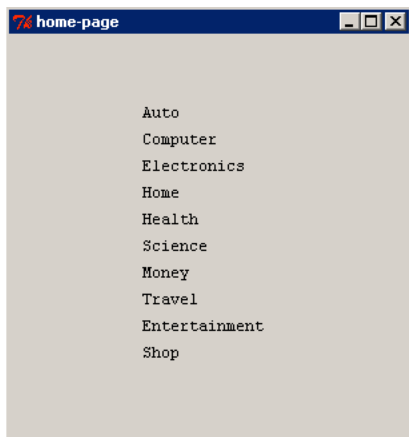


Figure 1. A example webpage presented to the model

For this demonstration, the model has the task to find out *how coffee works*, more simply stated, to find a webpage with information about *caffeine*. Therefore, the following goal has been set for our model:

ACT-R syntax	English
GOAL ISA COMPREHEND- HYPER-TEXT INTENTION "coffee" SELECTION NIL LAST-SELECTION NIL THRESHOLD NIL PAGE-SWITCH NIL STATE START	The goal chunk has the type comprehend hypertext, and the following slots: intention, selection, last selection, threshold, page switch, and state. Each slot has a value, for instance, the intention slot has the value "coffee".

The model listens to the options presented on a particular webpage and selects one that is goal relevant, that is, semantically similar with the word "Coffee". Semantic similarities are not computed by the model, but imported from a tool called Generalized Latent Semantic Analysis (GLSA; <http://glsa.parc.com>). Our model selects an option with a goal-relevance higher than a specified threshold; in this case the threshold is a free parameter and has the value 0.10. The following rule is used to select the goal relevant item:

ACT-R syntax	English
(P judge-goal-relevance =goal> ISA comprehend- hyper-text intention =intention selection nil state assess =retrieval> ISA similarity-fact term1 =intention term2 =term2 ==> =goal> selection =term2 state click)	If the goal is to comprehend hypertext, and a selection has not been made and the state of the goal is "assess", and a similarity fact has been retrieved then select the value of the second slot of the similarity fact and update the selection slot of the goal.

The option "shop" is selected because it was retrieved as having a similarity with coffee of .142 (higher than the threshold), and the shop page is displayed (Fig. 2).

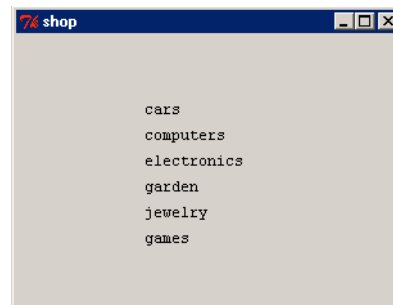


Figure 2. Example of a page loaded after a selection of the model. The option "shop" was selected on the previous page and the corresponding "shop-page" has been downloaded.

Starting with the second page, beside goal relevancies, information encountered on previous pages can be used when making current selections. Different ways to use past information are employed by the model and they will be discussed in separate sub-sections. The Annex presents a synopsis of information needed to understand the behavior of our model. It shows the most important Web pages, options on these Web pages (link labels), goal relevancies and several coherence coefficients. Goal relevancies are shown on the first line under each option. For example, the option "shop" on the home-page has a goal relevancy of 0.142. On the second line under some of the options, one can see a coefficient called "back-coherence1" – a semantic similarity between the current option and the previous selection of the model. For example, the option "garden" on the shop-page has a back-coherence1 of 0.33 (i.e., the semantic similarity between the current option "garden" and the model's past selection "shop" is 0.33). On the third line, a coefficient called "back-coherence2" can be seen. This is calculated as the semantic similarity between the current option and the model's past selection made 2 steps prior to the current step. For example, the option "supplies" on the garden-page has a back-coherence2 of 0.22 (i.e., the semantic similarity between the current option "supplies" and the model's selection at two steps back "shop" is 0.22). The coefficients back-coherence3 and back-coherence4 are defined in an analog way.

Forward linear navigation

For applying a selection strategy based only on goal relevancy, the model retrieves a goal relevancy higher than 0.10 using the following rule:

ACT-R syntax	English
(P retrieve-similarity =goal> ISA comprehend- hyper-text intention =intention state attend	If the goal is to comprehend hypertext, and the intention is temporarily stored in the variable "=intention", and the state of the goal is

<pre>?retrieval> state free ==> +retrieval> isa similarity-fact term1 =intention > value 0.1 :recently-retrieved nil =goal> threshold 0.1 state assess)</pre>	<pre>“attend”, and there is no other retrieval process currently performed, request a retrieval of a similarity fact having the intention (“coffee”) as the first term and a value higher than 0.1, that has not been retrieved recently, remember the threshold, and change the goal state in “assess”.</pre>
--	--

In case this retrieval attempt is successful, the model makes a selection and proceeds further on a new page. Based on this strategy, the model selects “shop” on the home-page and “garden” on the shop-page. On the garden page, no option is found with a goal relevancy higher than the threshold.

Backtracking and lowering the threshold

In case a goal relevancy cannot be found, the threshold is lowered (as in [18]) to 0.05 and the request for retrieval is repeated. If there are no goal relevant options on the current page, options encountered on previous pages are considered (as in [13]). ACT-R’s sub-symbolic computations (base-level activation in this case) ensure that previous options are considered in the inverse order of their perception (the more recent ones are considered before the older ones). In other words, the model backtracks one step at a time, as most of the users do by using the most common function of the back-button.

As a result of this strategy, the model deploys a very elaborate navigation behavior. From the garden-page in our example, it selects “supplies” then “pool”, it backtracks from the pool-page and selects “accessories” on the supplies-page, it backtracks again from the accessories-page and supplies-page to select “tools” on the garden-page, etc., etc. However, most of the selections it makes are not very helpful in understanding how coffee works. It finds some information about “hangover” and “alcohol” that is somehow related to “coffee”, but other more relevant options are not selected. In particular, the page containing the option “caffeine” is not selected, although the option “caffeine” itself has a high goal relevancy.

Coherence with past selections

Based on the previous strategy the model took too large detours and was trapped in too many ambiguities and ramifications of the natural language. For example, the word “shop” might have some semantic association with the word “coffee” but it has also thousands of associations with other words. When it jumps to “garden” (which is also somehow related to “coffee”), the initial association between “coffee” and “shop” is lost. By enforcing a coherence constraint one can ensure that the model stays within a limited set of meanings.

In case of this strategy, the model checks for both goal relevance and coherence with past selections. The model selects a goal relevant item only if its back-coherence1 is higher than 0.30. For this example, checking for back-

coherence1 was enough to make the model avoid highly inefficient divagations such as “decor”, “supplies” and “tools”. Checking for higher order coherences would have made the model far too conservative. While this strategy was good enough at improving model’s efficiency, it was not as good at improving the model’s effectiveness: pages relevant to how coffee works are still not selected. This time not even the “hangover” and “alcohol” options are selected.

Conservative and explorative strategies intertwined

The coherence strategy is useful to avoid divagations but also too conservative. It leads to blockages – situations where there is no goal relevant and coherent item to be selected and backtracking and lowering the threshold have already been applied. In these cases the model is allowed to employ the explorative strategy just until it overcomes the blockage. A rule that fires only in blockage situations omits the coherence check and selects based only on goal relevance. After this, the model goes back to the conservative (coherence-based) strategy. This combined strategy makes the model both efficient and effective (as effective as the explorative strategy can ensure). However, the model is still not effective enough: pages needed to understand, in our example, how coffee works are still not encountered.

Post-valued recall

So far, our model has only used current information displayed on the screen and information remembered from its own past selections. Information previously encountered but not selected has been ignored. However, users reassess previously discarded items based on newly acquired information, a phenomenon known as post-valued recall [27].

Whenever the model makes a selection of a link label to be clicked on, it also requests a retrieval of a previously non-selected item that has a relatively high semantic similarity (0.20) with the currently selected item. If such an item is found, it is selected. After such unusual selection the model continues its regular behavior – enforcing goal relevance and back coherence.

Based on this strategy the model found a high semantic similarity between “alcohol” and “drugs” (0.37), which led ultimately to finding the “caffeine” option on the addictive-page.

Discussion and conclusion

This model implements key features of Web navigation behavior as reported in literature and found in our empirical research. Some of these features are shared with previous models – selections based on goal relevance (information scent); backtracking, threshold and opportunistic strategies; back coherence – others are implemented here for the first time – intertwining between conservative and explorative strategies, and post-valued recall.

The ACT-R cognitive architecture allowed representation and simulation of the combination between syntactic and semantic aspects involved in Web navigation that has been shown to determine task performance. This syntax-semantics combination is inherent in the architecture. For example, a combination between “what” and “where”

allowed building a complete and functional memory representation of the information space being navigated. Based on this representation it was possible to combine semantic similarity judgments with syntactic backtracking strategies; an item is retrieved at a particular moment not only because of its content but also because of its position in the sequence of items that have been inspected: a recency effect modeled by the base-level activation of chunks simulated the order of items within a path – the most recent item has the closest position to the current item; and spreading activation ensured retrieval of the chunk required by a particular state of the environment – a memory chunk is more likely to be retrieved if it is associated with the goal chunk.

There is still further work needed to ensure full cognitive plausibility and fit of the model to human data. We claim that our modeling decisions are founded in literature and in our own empirical research, but we also admit that more empirical and modeling efforts are needed to fine tune some of the parameters of the model. For example, the existence of a selection threshold and the satisficing behavior are generally accepted [17,18]; however we don't know the exact value of this threshold and to what extent it should be lowered during backtracking. The values we have used (0.1 and 0.05) are based on our experience with GLSA-like tools and modeling and experimentation with a variety of Web navigation tasks and websites. The same kind of "educated guesses" have been used for the back coherence and post-valued recall parameters.

In conclusion, in spite of the aforementioned limitations, the model presented here demonstrated a way to proceed toward building a fully computational cognitive model of Web navigation. This model needs to be extended in order to gain more automation and autonomy, that is, it should be less dependent on user input and able to interact with a standard interface and with other models/agents. Significant work is also needed for making this kind of model scalable to large information spaces.

GENERAL CONCLUSION AND DISCUSSION

This paper has addressed some theoretical and methodological aspects of the problem of using the Internet by VIPs via screen readers. The empirical study explored behavioral aspects involved in using the Internet with the standard screen turned off and, instead, with the aid of a screen reader. Some of the results were similar with results found in previous studies, showing that common cognitive mechanisms operate in both contexts of use. In particular, spatial ability was found to be an important determinant of web task performance, even in this particular context of use in which all the visual cues are absent. Other results were specific to using the web via screen reader. For example, visiting a high number of pages has the potential to cause disorientation in VIPs, whereas for sighted users it has a opposite effect, as we have found in a previous study [10].

In addition, the empirical study tested the effectiveness of a particular manner of providing model-based navigation support. The particular technique we used – increasing the reading priority of goal relevant items – was chosen because it did not create information overload. However,

the effectiveness of this technique could not be proven. A possible reason for this failure could be that this technique altered the inherent coherence of content on web pages. The lesson learned from this failure was that providing model-based navigation support for VIPs should not break the intrinsic coherence of the content on Web pages.

Development of a computational model of using the web via screen readers is possible and opportune. This paper presented only a demonstration of such model and a first attempt at what seems to be a non-trivial problem. The model needs further development and comparison with human data. A future goal worth pursuing would be to integrate this type of model in a more complex tool supporting the web use of VIPs, that is, assisting the user in making goal relevant selections and avoiding part of the redundancy associated with the use of screen readers.

REFERENCES

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036-1060.
2. Bothell, D. (2005). ACT-R 6 Official Release. *Twelfth Annual ACT-R Workshop*, Trieste, Italy: Universita degli Studi di Trieste.
3. Brumby, D. P., & Howes, A. (2007). Strategies for Guiding Interactive Search: An Empirical Investigation into the Consequences of Label Relevance for Assessment and Selection. *Human-Computer Interaction*.
4. Budiu, R., & Pirolli, P. (2007). Modeling Navigation in Degree-of-Interest Trees. *Paper presented at the CogSci*, Nashville, TN.
5. Cockburn, A., Greenberg, S., Jones, S., McKenzie, B. J., & Moyle, M. (2003). Improving WEB Page Revisitation: Analysis, Design and Evaluation. *IT & Society*, 3(1), 159-183.
6. Di Blas, N., Paolini, P., & Speroni, M. (2004). "Usable accessibility" to the Web for blind Users. *Proceedings UI4ALL*, Vienna.
7. Fauemberger, C., Stockman, T., Putz, V., & Hoeldrich, R. (2005). Interaction Patterns for Auditory User Interfaces. *Paper presented at the International Conference on Auditory Display*, Limerick, Ireland.
8. 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.
9. Goldman, S. R., Golden, R. M., & Van den Broek, P. (2006). Why are Computational Models of Text Comprehension Useful? in F. Schmalhofer & C. A. Perfetti (Eds.) *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ.: L. Erlbaum Associates.
10. Herder, E. & Juvina, I. (2004). Discovery of Individual User Navigation Styles. *Adaptive Hypermedia AH2004*, Eindhoven.
11. Hurst, W., Lauer, T., Burfent, C., & Gotz, G. (2005). *Forward and Backward Speech Skimming with the Elastic Audio Slider*. Paper presented at the HCI'05

- Invited paper presented at RoCHI 2007, Constanta, Romania.
- Conference on People and Computers XIX, Edinburgh, UK.
12. 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. *Proceedings of Human Factors and Ergonomics Society 46th Annual Meeting*. Pittsburgh, PA, USA.
 13. Juvina, I. & van Oostendorp, H. (2005). Cognitive Model Working Alongside the User. *British HCI2005*, Edinburgh, UK.
 14. Juvina, I., & van Oostendorp, H. (2006a). Individual differences and behavioral metrics involved in modeling web navigation. *Universal Access in Information Society*, 4, 258-269.
 15. Juvina, I., & van Oostendorp, H. (in press). Modeling Semantic and Structural Knowledge in Web Navigation. *Discourse Processes*.
 16. 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*, 357-373.
 17. Krug, S. (2000). Don't Make Me Think! A Common Sense Approach to Web Usability, Indianapolis, Indiana: New Riders Publishing.
 18. Miller, C. S., & Remington, R. W. (2004). Modelling Information Navigation: Implications for Information Architecture. *Human-Computer Interaction*, 19(3), 225-271.
 19. Morrison, J. B., Pirolli, P., & Card, S.K. (2001). A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. *CHI 2001*, Seattle.
 20. Neerinx, M.A., Lindenberg, J. & Pemberton, S. (2001). Support concepts for Web navigation: a cognitive engineering approach. *Proceedings Tenth World Wide Web Conference*, pp. 119-128. New York, NY: ACM Press.
 21. 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. *Proceedings of the Eighth ACM Conference on Hypertext*. Southampton, UK.
 22. 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: Springer-Verlag.
 23. 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.
 24. Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
 25. Strain, P. & McAllister, G., Enhancing Web Accessibility for the Visually Impaired using RSS, *Paper presented at the HCI'05 Conference on People and Computers XIX*, Edinburgh, UK.
 26. 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. Zeist, The Netherlands: Bartimeus Instituut.
 27. Wen, J. (2003). Post-Valued Recall Web Pages: User Disorientation Hits the Big Time. *IT&Society*, 1(3), 184-194.
 28. Yu, W., McAllister, G., Strain, P., Kuber, R. and Murphy, E., Improving Web Accessibility Using Content-Aware Plug-ins, *Paper presented at the HCI'05 Conference on People and Computers XIX*, Edinburgh, UK.
 29. Zajicek, M., Powell, C., & Reeves, C. (1998). Orientation of Blind Users on The World Wide Web. M. Hanson (ed.) *Contemporary Ergonomics*. Taylor and Francis, London.

Annex

A simplified version of the website How Stuff Works together with coefficients used by the ACT-R model presented above. The user intention is to find out how coffee works. The numbers presented on the first line below each option are goal relevancies – semantic similarities between the word “coffee” and each link label, respectively. Numbers appearing on the second line are called back-coherence1 coefficients – semantic similarities between the previous selection and each option, respectively. Back-coherence2 coefficients are shown on the third line and they are calculated as semantic similarities between a two-steps back selection and each option, respectively. Back-coherence3 and back-coherence4 coefficients are calculated in an analog way.

Home-page

auto	computer	electronics	Home	health
.002	.007	.008	.039	.088
science	money	travel	entertainment	shop
.069	.038	.000	.079	.142

Shop-page

cars	computers	electronics	garden	jewelry	games
.044	.007	.008	.125	.049	.008
.44	.20	.08	.33	.14	.29

Health-page

diseases	drugs	Fitness	care	nutrition	pregnancy
.015	.030	.009	.056	.010	.011
.48	.47	.06	.59	.19	.09

Science-page

engineering	life	Supernatural
.009	.043	.019
.44	.24	.02

Entertainment-page

arts	games	movies	music	sports	television	toys
.038	.008	.001	.005	.006	.004	.116
.11	.24	.36	.26	.29	.44	.19

Drugs-page

Care-page

performance	Enhancing	nicotine	diet	hangover
.034	.018	.129	.036	.051
.02	.00	.03	.12	.08
.02	.04	.01	.27	.04

Toys-page

Garden-page

doodle	radio		décor	supplies	tools	irrigation
.079	.027		.052	.104	.069	.011
.15	.02		.05	.13	.01	.06
.01	.22		.09	.22	.23	.02

Tools-page

mowers	shears	clippers
.045	.004	.008
.06	.01	.06
.06	.03	.01
.00	.02	.02

Nicotine-page

brain	addictive	withdrawal	toxicity	cancer
.027	.064	.049	.044	.029
.04	.35	.12	.05	.00
.31	.02	.05	.01	.38
.20	.01	.02	.02	.41

Hangover-page

Doodle-page

alcohol	nausea	aspirin		pen	eraser
.082	.011	.052		.015	.024
.19	.06	.08		.05	.19
.13	.08	.05		.03	.03
.20	.05	.05		.00	.02

Decor-page

fountains	plaques	Vanes
.024	.014	.006
.07	.01	.01
.04	.06	.03
.00	.01	.05

Supplies-page

pool	accessories	Filters	pumps
.126	.067	.028	.039
.08	.21	.13	.21
.19	.04	.03	.04
.10	.21	.04	.05

Addictive-page

Alcohol-page

caffeine	marijuana		effects	warning	abuse
.147	.023		.003	.045	.029
.24	.10		.01	.01	.35
.28	.17		.04	.01	.09
.09	.22		.19	.10	.20
.05	.01		.29	.10	.18

Aspirin-page

Pool-page

headache	heart		Covers	heaters	filters	pumps
.010	.014		.024	.021	.028	.039
.03	.03		.00	.01	.01	.04
.13	.01		.24	.00	.13	.21
.01	.43		.14	.03	.03	.04
.03	.19		.18	.04	.04	.05

Accessories-page

fence	spa
.014	.019
.06	.03
.05	.04
.05	.03
.03	.07