InfoScent Evaluator: A semi-automated tool to evaluate semantic appropriateness of hyperlinks in a web site

Katsanos Christos*, Tselios Nikolaos*#, Avouris Nikolaos*

*Human-Computer Interaction Group, Electrical and Computer Engineering Dept., University of Patras, Rio, Greece

#ICT in Education Group, Dep. of Educational Sciences and Early Childhood Education University of Patras, Rio, Greece

{ckatsanos, nitse, avouris}@ece.upatras.gr

ABSTRACT

In this paper, we present InfoScent Evaluator, a tool that automatically evaluates the semantic appropriateness of the descriptions of hyperlinks in web pages. The tool is based on a theoretical model of users' behavior when engaged in information search tasks, called Information Foraging. A textual description of the user's search goal is compared with the textual description of each probable hyperlink, using Latent Semantic Analysis, a statistical technique that evaluates the distance between the two texts. Through this approach the most probable path that the user will follow in order to access the sought web page can be predicted. Thus, the tool can be used to evaluate the web site in terms of appropriateness of hyperlink text and of information architecture. We argue that the presented tool could substantially aid design and evaluation of a web site.

Author Keywords

Information Foraging theory, proximal cue, web usability evaluation, semi-automated tools, LSA index

ACM Classification Keywords

H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces – Evaluation/Methodology, H5.4 Information interfaces and presentation (e.g., HCI): Hypertext/Hypermedia – Navigation, Architectures, Theory, User Issues.

INTRODUCTION

During the last years, an exponential growth in the number of available web sites and applications is observed. However, proper design of such hypertext systems is not a straightforward process and still remains immature in a number of issues. Additionally to this, despite the proliferation of the Human Computer Interaction field, the research area still suffers from lack of sound theoretical descriptive or prescriptive models. Such a promising model is based on the notion of information foraging. The core idea of this theory is that information foraging is analogous to the food foraging mechanisms, therefore models of optimal foraging theory developed by anthropologists and ecologists in the study of food foraging will help us understand foraging

behavior in consumers of information.

In the initial work by Pirolli and Card (1995) on information foraging, they defined the profitability of an information source as "the value of information gained per unit cost of processing the source." Cost is defined in terms of time spent, resources utilized and opportunities that are lost when pursuing one particular strategy instead of others. Furnas (1997), coined the term 'residue' to describe the hint that a representational object holds (e.g. a hyperlink) of what lays behind it. Therefore information foraging theory could be used as a tool to examine and model user interaction with Web sites. In particular, it could be treated as a way to examine user goals, their decision making processes and adaptations to the information access system environment. Most importantly, information foraging theory can lead both the design and the evaluation process. Researchers can then make use of this knowledge in assessing system and interface design.

The task of evaluating and improving the usability of websites can be daunting given the quantity of sites being produced, the frequency of updates and the sheer size of many sites (Brinck and Hofer, 2002). Additionally to proper and deep understanding of the cognitive processes taking place during search of information in a web site, web designers, usability specialists and researchers need tools that permit the rapid exploration of hypotheses about complex interactions of user goals, user behaviours and Web Site designs. Especially for practitioners, it is crucial to offer an increased level of automation in the usability evaluation process, since they function under strict time constraints (Chi et al., 2000). We argue that a tool which semi-automates the evaluation of appropriateness of links' descriptions is of fundamental importance, since it seems to be a strong determinant of users' satisfaction, even more than proper content organisation. In a relative study, Resnick and Sanchez (2004) showed that for sites with good labels and links' descriptions there was no benefit to having one site structure over the other.

In the research presented in this paper, we attempt to tackle this issue. We present a tool which automatically evaluates semantic appropriateness of the hyperlinks' descriptions. The tool is based upon the notion of Information Foraging theory and attempts to quantify the concept of information *scent* using a statistical technique, named Latent Semantic Analysis (LSA). The tool can be used to evaluate a web site for a variety of issues, as described in the following.

The rest of the paper is organized as follows: First, we present a brief overview of Information Foraging Theory and Latent Semantic Analysis. Subsequently, we present the architecture of the tool realised to aid the evaluation of links' appropriateness of a web site, named InfoScent Evaluator (ISEtool, freely available at hci.ece.upatras.gr). Additionally to this, we present our vision for the method which is realised with the existence of ISEtool. Description of the tool is provided through a detailed example of its application.

INFORMATION FORAGING THEORY

Information foraging theory is grounded in computational theories of human cognition and optimal foraging theories from biology. According to the theory of information foraging users assess the appropriateness of following a particular path on the Web by considering a representation, usually a textual description or graphic, of the distal content. Residue was recast and refined by Pirolli (1997) as information "scent" and defined in Card et al. (2001) as a user's "(imperfect) perception of the value, cost, or access path of information sources obtained from proximal cues, such as WWW links".

LATENT SEMANTIC ANALYSIS

LSA (Landauer & Dumais, 1997) was developed to mimic human ability to detect deeper semantic associations among words, phrases or whole sentences. The latter technique served as a computational model of information scent in other scientific efforts, such as ACWW, a conceptual artifact based on Cognitive Walkthrough Evaluation method and a theory analogous to Information Foraging, named Colides (Blackmon et al., 2005). However, despite the fact that the results presented seem to be very promising, lack of integration of useful functions - such as automatic grabbing of links, storage of results, automatic prediction and walkthrough of a user's path) - inspired us to create a more complete and automated tool described in detail in the following.

INFOSCENT EVALUATOR

The main components of the ISEtool architecture are diagrammatically represented in Figure 1. The basic underlying assumption in our tool is that human information-seeking and sense-making behavior is explained and modeled sufficiently by the Information Foraging Theory (Pirolli, 1999). Thus, the theoretical model we adopt is that users have some information goal and their surfing patterns through the site are guided by information scent. LSA is used to exploit computationally the theoretical model by quantifying this concept of information scent.

A typical usage scenario of the tool is the following: First, the Evaluator describes a typical user goal in a text box using free text. Subsequently, the ISEtool parses the main page of the Site under examination. More at the point, it collects all available links on the page, it finds and stores link kind (e.g. image or text?) and type of the pointed file (e.g. extension). It also finds and stores the "proximal cue" of each link by grabbing the textual description, if it is a text hyperlink or the alternative text

(e.g. ALT tag), if it is a graphical hyperlink. The calculation of "Information Scent" for all links is achieved by running in a transparent and automated way one-to-many analysis of the LSA algorithm (http://lsa.colorado.edu/). LSA computes and returns the semantic similarity (LSA index) of the user goal against all the proximal cues of the links. Additionally to this, the tool also discriminates automatically the external and internal links. Next, the tool offers a number of options (described in more detail in the representative example section) to the Evaluator. For example, it can automatically select the link with the higher LSA index and subsequently "visit" the next page by repeating the same process until the user goal is met or there is no "good" link for this goal.

The tool can be used to identify a variety of navigability and usability problems. First, at a page level it could identify "Competing Links". When two or more link-labels have comparable LSA index, then possibly this can cause confusion to the user. S/He has to decide which path to take, since more than one seems to have comparable probability to lead to the information goal.
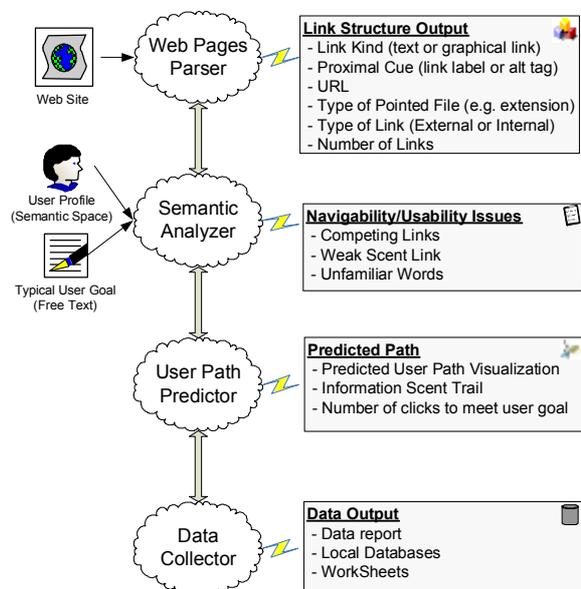

Figure 1. Architecture of the ISEtool

In addition, it can identify pages containing "Weak Scent Links", which can happen when all labels have a relative low (below a user defined threshold) LSA index. A page could be also checked for cognitive overload issues, if it contains a plethora of hyperlinks possibly overwhelming the user. Moreover, the tool can provide strong indications of "Unfamiliar Words" by taking advantage of the LSA term vector length, a measure that is correlated with word frequency. When some of the words used in the proximal cues of the links have short term vector lengths (e.g. they have low frequency in the defined corpus) this is a good indication that the users modeled by the semantic space selected will perceive them to be relatively meaningless and thus they will be incapable of relating them semantically to their goal.

Subsequently, at a *site level*, and by using the visualization of the predicted path offered by the tool, the
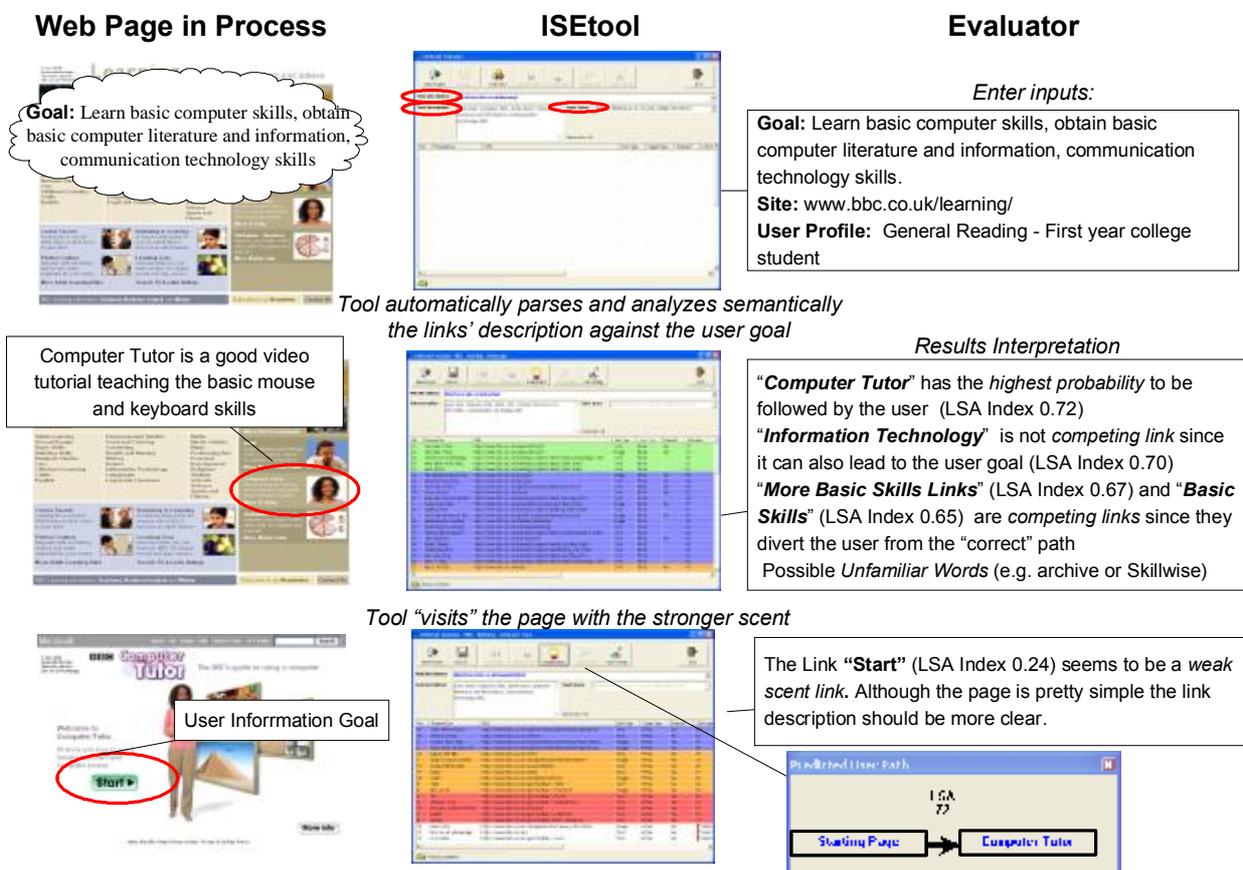
Figure 2. A representative example of using InfoScent Evaluator Tool to analyze the BBC learning home page (http://www.bbc.co.uk/learning/) against a user information goal

researcher can see at which page the user could possibly quit the site since the scent falls below a specific threshold. Additionally to this, the length of the predicted path to reach the desired information goal can be calculated. Research has shown that the likelihood for a user to get frustrated is significantly increased after 3 clicks (Huberman et al., 1998). Therefore, lengthier paths are an indication of poor information architecture, which is a source of various navigability and usability problems.

A REPRESENTATIVE EXAMPLE

In order to present to a greater extent the utility of ISEtool we demonstrate its use by a representative example, shown in Figure 2. In this example, the Evaluator defines the URL of the Site under examination, a typical user information goal and a typical user profile (e.g. LSA semantic space). Subsequently, the tool first parses some useful attributes of the hyperlinks in the starting page (e.g. links' description, URL, etc) and then analyzes the semantic similarity of all the links' descriptions against the user goal. At this point, the Evaluator inspects the link structure produced by the tool. In order to get a better insight s/he decides to sort the links by their semantic appropriateness (LSA index) and change the default color coding (e.g. define ranges and colors of semantic similarity) adjusting the tool to his/her needs and preferences. Subsequently, s/he realizes that there are some proximal cues with comparable LSA index that could possibly confuse the user (e.g. "Computer Tutor" –

0.72, "Information Technology" – 0.70, "More Basic Skills Links" – 0.67, "Basic Skills" – 0.65). By examining the targeted pages of these links s/he quickly realizes that the first two links lead to the user goal and thus they are not classified as competing links. On the other hand the latter two links divert the user from the "correct" path and thus they are classified as competing links.

At a next step, the evaluator decides to use the tool to automatically select the link found to "emit" the higher scent (e.g. "Computer Tutor") and subsequently "visit" the next page by repeating the same process until the user goal is met or there is no "good" link for this goal. Furthermore, s/he chooses to export the analyzed pages to a worksheet in order to elaborate them in different ways.

It is worth mentioning that at any step of the simulation and at a site level analysis the evaluator is offered a visualization of the predicted path and the scent trail followed. Thus, s/he can identify cases of pages having low scent (e.g. below a threshold that s/he defines) that could lead the user to leave the site, and investigate them further.

By taking advantage of this battery of functions the user of the ISEtool can investigate navigability and usability problems (like competing or weak scent links) or problems that relate to cognitive overload issues (e.g. the total number of links). Additionally, the tool generates automatically some warnings of possible unfamiliar

words for the user profile selected (e.g. word "archive"). These warnings could be translated to strong indications that the user will probably find meaningless some words and thus will be unable to relate them semantically with his/her goal.

CONCLUSION

This paper presented InfoScent Evaluator tool (ISEtool) which automatically evaluates semantic appropriateness of the hyperlinks' descriptions of a web site. The tool is based upon the notion of Information Foraging theory and attempts to quantify the information scent, using a statistical technique, named Latent Semantic Analysis (LSA). The utility of the tool is twofold. First, during the early phases of design, it allows Web-Site designers to explore alternative designs and solutions ("what-if" scenarios) but also supports usability specialists in conducting formative usability evaluations, reducing the need to involve actual users which is costly and time-consuming. Initial application of the tool showed promising results.

However, at the moment, the presented tool has some limitations. For instance, initially when the web page is analyzed, the alternative text (e.g. the ALT Tag) of each graphical-link is assumed to be the proximal cue of this link. Nevertheless, even if all the graphical-links have alternative text - which is usually not the case - this approach is still questionable, since for a user with a graphical browser the actual proximal cue is the graphic itself and not the alternative text. The way provided to overcome this problem, but also to allow the definition of proximal cues for graphical-links with no alternative text, is to notify the user of the tool for such instances and ask for his/her descriptions. Additionally to this, some web sites present a set of links in a nested menu-like approach. In such cases, the user primarily, focuses on the general description and spreads the rest of his/her attention to identify the desired menu subitem, which is anticipated by taking into account the context communicated by the header. At the moment, our tool cannot mimic this behaviour, thus often leading to 'flat' or misleading results. Furthermore, the automatic warnings generated as an indication of unfamiliar words are not always trustworthy due to the fact that LSA semantic spaces are produced using a corpus "of formal documents" and thus "informal" language or abbreviations - that are often used in web pages – may lead to unexpected results.

Future work contains conducting studies with real users, in order to compare the obtained results with tool results and further validate the accuracy of the developed method/tool. Additionally to this, to study how goal reformulation happens when users browse a web site towards finding desired patches of information. Finally, to a broader context, to examine possible extensions to the developed model (e.g. parameterize the infoscent model with data describing user profiles or specific particularities of the information domain of the specific

information to be found). Further research is needed in order to investigate these areas, together with longitudinal evaluation of the effects of other aspects of a web site, such as credibility and aesthetics (Papachristos et al., 2005).

REFERENCES

Blackmon, M, Kitajima, M. & Polson, P. (2005). Tool for Accurately Predicting Website Navigation Problems, Non-Problems, Problem Severity, and Effectiveness of Repairs. Proc. CHI'05, ACM Press, pp.31-40.

Brinck, T. and Hofer, E. (2002). Automatically evaluating the usability of web sites. Workshop Organizer at CHI'02.

Card, S., Pirolli, P., Wege, M., Morrison J., Reeder, R. W., Schraedley, P, Boshart, J. (2001) Information scent as a driver of Web behavior graphs. Proc CHI '01, ACM Press, 498-505.

Chi, E., Pirolli, P., & Pitkow, J. (2000). The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a website. Proc. CHI 2000, ACM Press, 161–168.

Furnas, G. (1997). Effective view navigation. Proc. CHI'97, ACM Press, 367–374.

Huberman, B., Pirolli, P., Pitkow, J. and Lukose, R. (1998). Strong Regularities in World Wide Web Surfing. Science, 280, 95-97.

Landauer, T. & Dumais, S. (1997) A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 211–240.

Papachristos, E., Tselios, N., Avouris, N. (2005) Inferring relations between color and emotional dimensions of a web site using Bayesian Networks. Interact 2005, Lecture Notes in Computer Science, Volume 3585, Sep 2005, Pages 1075 - 1078..

Pirolli, P. & Card, S. (1995) Information Foraging in Information Access Environments. Proc. CHI '95. ACM Press/Addison-Wesley Publishing Co, 51-58.

Pirolli, P. (1997). Computational models of information scentfollowing in a very large browsable text collection. Proc. CHI '97, ACM Press, 3-10

Pirolli, P., & Card, S. (1999) Information foraging. Psychological Review, 106, 643–675.

Resnick, M. L. & Sanchez, J. (2004). Effects of organizational scheme and labeling on task performance in product-centered and user-centered retail websites. Human Factors, 46.