History-based Link Annotation for Self-Exploratory Learning in Web-based Hypermedia

Muan H Ng, Wendy Hall, Patricia J Maier

University of Southampton, UK

E-mail: mhn99r, wh, pjm00r@ecs.soton.ac.uk

Ray Armstrong

Southampton General Hospital, UK

E-mail: rayarmstrong@btinternet.com

ABSTRACT

In this paper, we examine the notion of history-based link annotation for self-directed exploratory learning environment. We first explain the motivation for exploratory learning in web-based hypermedia and investigate the effectiveness of adaptation to reduce cognitive overload. We then present a simple user model that considers individual reading speed, comprehension rate and prior knowledge to determine effort spent on a page. An initial pilot study is conducted to support the work described in this paper.

KEYWORDS: Adaptive navigational support, history-based link annotation, reading speed.

INTRODUCTION

Browsing is a form of informal learning in hypermedia [1] and it is intrinsic to the nature of hypermedia – a tool to obtain information. As opposed to formal learning where a set of rigid paths and tasks are exposed, users are given the freedom to explore information stored in the hypermedia. The nature of browsing is driven more by the curiosity and presence of anchors rather than a definite learning goal. As users navigate from one node to another, pieces of information bits are accumulated to construct knowledge blocks. A *history-based link annotation* is proposed to aid cognitive overload in this style of learning. Links are annotated based on the users' browsing history to inform them which links they have already visited. Ideally, users can decide to pay more attention to pages that have been 'less attended to' from previous browsing.

PREVIOUS WORK

The system first needs to determine whether a user has indeed visited a page. Previous work in Interbook[2] recorded the display of a page to determine if the page has been read. But this method lacks accuracy as a student might only be surfing through the pages without paying much attention to the content. In MANIC[3], the system attempts to determine if the students has in fact read a page by considering time spent. An optimal reading time based on the length of a page is used to compare with the actual time spent. It assumes a high studied rating for a page if user spends an optimal amount of time on it. However, the optimal time that is generated based on the content displayed is static for every user irrespective of their individual reading speed. As far as we know, none of the existing adaptive hypermedia application considers individual differences in determining whether a page is read.

USER MODEL: CAPTURING EFFORT SPENT ON CONTENTS

Literature in psychology has proved that there are "astonishing differences in the rate of reading speed" among individuals [4]. This applies even if they are reading at normal speed. It has also been found that two individuals reading at the same speed can have different comprehension rates [5]. Similarly, if a subject reads comparatively slowly, they do not necessarily comprehend more than a fast reader does. This work has proved that both the **speed** and **power of assimilation** can affect reading efficiency and there is no direct relationship between them.

On top of these, individual reading style can also affect the reading speed. A person who knows the domain well may tend to read at a rapid rate to review familiar materials. By contrast, a person who is a novice may read at a normal rate to grasp relations to general ideas. Hence we consider another factor of '**prior knowledge**' to make a more accurate prediction of the user's reading speed. Users have to perform a self-rating on the domain knowledge for the system to obtain a value for prior knowledge.

We have built our user model based on the three individual factors formulating the notion of effective reading speed. This notion is greatly influenced by standard ways of measuring reading speed in commercial software solutions to enhance reading [6]. In our case, a factor of prior knowledge is added to refine the existing standard.

Hence, we defined the effective reading speed as the actual reading speed of a user weighted by his/her comprehension rate and prior knowledge.

$$\textit{Effective reading speed} = \textit{reading speed} \times (\textit{comprehension rate} + \textit{prior knowledge}) / 2$$

For example, assuming a reader who can read at the speed of 200 words per minute. If he/she can comprehend 60% of the materials read (derived from the percentage of correct answers in a post-reading test), and his/her self-assessment of prior knowledge is 70%, the *effective* reading speed becomes 200 x (0.6+0.7) / 2, which is equivalent to 130 effective words per minute (ewpm).

HISTORY-BASED LINK ANNOTATION

In the process of browsing and navigating, the individual effective reading speed is used as a benchmark to build a history of browsing for each individual. An *optimal* time spent for each page in the domain is generated dynamically based on the length of the page and the individual effective reading speed:

Optimal time for a page, T
= number of words / effective reading speed

When a user spends an amount of time, t, on a page, we can determine the effort spent on that page by comparing t with optimal time, T. A page is considered well studied if the actual time spent approaches the optimal time ($t \approx T$). If the user spends too much or too little time on the page we assume that he/she is skipping the page or has left the application unattended. In that case, a zero effort is assumed (see Figure 1).

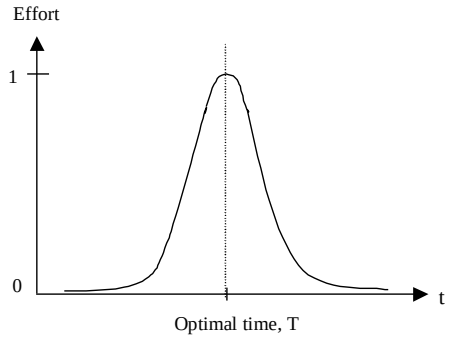

Figure 1: A normal distribution of effort based on time spent

The history of browsing is then used to annotate links as in Figure 2. A zero percentage means the user has never or has hardly looked at the information associated with Link8 and a 100% means the user has spent enough "effort" for the system to assume that the information associated with Link3 is well read.

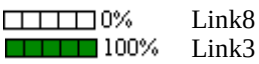

Figure 2: Links annotated by the percentage of effort

PILOT STUDY

A pilot study has been conducted to compare an application with and without the history-based (HB) link annotation. The subjects are tested on a web-based application for the study of Rheumatology. Initial results show that 63% out of 19 subjects found the HB link annotation useful. The results also show that the average time spent on the case with the HB link annotation is reduced for subjects who are novice to the domain. However the effect of the HB link annotation is not significant for subjects who are expert in the domain. In observing the number of links clicked, comparatively fewer links are followed in the case with the HB link annotation, provided if the list of the links are long (>10 links).

CONCLUSIONS

Our conclusion so far is that the HB link annotation might help to reduce cognitive overload in self-exploratory learning. We have demonstrated the use of individual reading speeds to model users' browsing history. However, even though the system only approximates the 'effort spent' on a page, to a certain extent, the 'students' understanding' of the information read is estimated since we take into consideration their comprehension rate and prior knowledge. Further work and evaluation will be done to find any significant learning effects, given the integration of the history-based link annotation into the hypermedia environment.

ACKNOWLEDGMENTS

We would like to thank the Arthritis Research Campaign for funding the project.

REFERENCES

1. Duchastel P. Formal and Informal learning with hypermedia, in Designing Hypermedia for Learning, Springer-Verlag, New York, 1989.

2. Brusilovsky P. and Schwarz E. User as Student: Towards an Adaptive Interface for Advanced Web-Based Application. In Proc of UM97. Springer-Verlag, New York (1997), pp 177-188.

3. Stern M. and Woolf B.P. Adaptive Content in an Online Lecture System, in Proc of the AH 2000, Trento, Italy (2000), pp 227-238

4. Romanes G. J. Mental Evolution in Animals, Appleton & Co, New York, 1885.

5. Huey E. B. The psychology and pedagogy of reading, The Macmillan Company, 1908.ACM Press, pp. 381-386.

6. ReadingSoft.com found at http://www.readingsoft.com