

SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web

Wai-Tat Fu (wfu@uiuc.edu)
Human Factors Division and Beckman Institute
University of Illinois at Urbana-Champaign
405 North Mathews Avenue
Urbana, IL 61801, USA

Peter Pirolli (pirolli@parc.com)
Palo Alto Research Center
3333 Coyote Hill Rd
Palo Alto, CA 94304, USA

Manuscript submitted to Human-Computer Interaction
Date: Jan 03, 2007

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 30 JAN 2007		2. REPORT TYPE		3. DATES COVERED 00-00-2007 to 00-00-2007	
4. TITLE AND SUBTITLE SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Illinois at Urbana-Champaign, Human Factors Division and Beckman Institute, 405 North Mathews Avenue, Urbana, IL, 61801				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 67	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

ABSTRACT

We describe the development of a computational cognitive model that explains navigation behavior on the World Wide Web (WWW). The model, called SNIF-ACT (Scent-based Navigation and Information Foraging in the ACT cognitive architecture), is motivated by Information Foraging Theory (IFT), which quantifies the perceived relevance of a Web link to a user's goal by a spreading activation mechanism. The model assumes that users evaluate links on a Web page sequentially, and decide to click on a link or to go back to the previous page by a Bayesian satisficing model (BSM) that adaptively evaluates and selects actions based on a combination of previous and current assessments of the relevance of link texts to information goals. The model was tested against data collected from novice users engaged in unfamiliar information-seeking tasks. SNIF-ACT 1.0 utilizes the measure of utility, called *information scent*, derived from IFT to predict rankings of links on different Web pages. The model was tested against a detailed set of protocol data collected from eight subjects as they engaged in two information-seeking tasks using the WWW. The model provided a good match to subjects' link selections and decisions to leave a Web site, and thus provided support for the use of information scent as a psychological measure of the perceived relevance of link text to information goals. In SNIF-ACT 2.0, we include an adaptive link selection mechanism that sequentially evaluates links on a Web page according to their position. The mechanism was derived based on a rational analysis of link selection on a Web page. The mechanism allowed the model to dynamically update the evaluation of actions (e.g., to follow a link or leave a Web site) based on sequential assessments of link texts on a Web page, and to decide when to leave a page based on experiences with previous pages. SNIF-ACT 2.0 was validated on a data set obtained from 74 subjects. Monte Carlo simulations of the model showed that SNIF-ACT 2.0 provided better fits to human data than SNIF-ACT 1.0 and a Position model that used position of links on a Web page to decide which link to select. We conclude that the combination of the IFT and the BSM provides a good description of user-Web interaction. Practical implications of the model are discussed.

1. INTRODUCTION

Most everyday problems, such as making an investment, planning travel around traffic conditions, or finding a restaurant, are ill-defined (Reitman, 1964; Simon, 1973) and require additional knowledge search (Newell, 1990) in order to develop a solution. A substantial number of people now turn to the World Wide Web in search of such knowledge.¹ Consequently, the Web has become a domain that allows the study of complex everyday human cognition. The purpose of this article is to present a computational cognitive model that simulates how people seek information on the Web. This model is called SNIF-ACT, which stands for Scent-based Navigation and Information Foraging in the ACT architecture. SNIF-ACT provides an account of how people use *information scent* cues, such as the text associated with Web links, in order to make navigation decisions such as judging where to go next on the Web, or when to give up on a particular path of knowledge search. SNIF-ACT is shaped by rational analyses of the Web developed by combining the Bayesian satisficing model (Fu & Gray, 2006; Fu, in press) with the information foraging theory (Pirolli, 2005; Pirolli & Card, 1999), and is implemented in a modified version of the ACT-R cognitive architecture (Anderson et. al, 2004)². In this article, we will describe the current status of the SNIF-ACT model and the results from testing the model against two data sets from real-world human subjects. At this point, our goal is to validate the model's predictions on unfamiliar information-seeking tasks for general users. To preview our results, our model was successful in predicting users' behavior in these tasks, especially in identifying the "attractor" pages that most users visited.

This paper reports on two versions of SNIF-ACT (versions 1.0 and 2.0) that have been developed to model how users navigate through the Web in search of answers to specific information-seeking tasks. SNIF-ACT 1.0 (Pirolli & Fu, 2003) was developed to simulate a small number of users working on a small number of tasks, whose Web navigation behavior had been previously subjected to very detailed protocol analysis (Card *et al.*, 2001). SNIF-ACT 1.0 establishes how information scent is used in navigation, but makes the strong assumption that all links from a Web page are attended and assessed prior to a decision about the next navigation action. SNIF-ACT 2.0 extends the first version of the model by incorporating the Bayesian satisficing model (Fu & Gray, 2006; Fu, in press) in the evaluation of Web links. The process of satisficing assumes that, instead of searching for the optimal choice, choices are often made once they are good enough based on some estimation of the characteristics of the environment. We also show that the user data and SNIF-ACT 2.0 Monte Carlo data can both be fit by the *Law of Surfing* (Huberman *et al.*, 1998), a

¹ Internet use is estimated to be 68.3% of the North American population (<http://www.internetworldstats.com/stats.htm>). It is estimated that 88% of online Americans involve the Internet in their daily activities (http://www.pewinternet.org/pdfs/PIP_Internet_and_Daily_Life.pdf).

² We modified the utility calculations of productions in the original ACT-R by a new set of calculations presented in later sections.

strong empirical regularity describing the distribution of lengths of navigation paths taken by users before giving up.

One reason for developing SNIF-ACT is to further a psychological theory of human information foraging (Pirolli & Card, 1999) in a real world domain. Real world problems pose productive challenges for science. New theory often emerges from scientific problems that reflect real phenomena in the world. Such theories are also likely to have implications for real problems that need to be solved. Psychological models such as SNIF-ACT are expected to provide the theoretical foundations for cognitive engineering models and techniques of Web usability. Following our presentation of SNIF-ACT, we discuss the relation of the model to a semi-automated Web usability analysis system called Bloodhound (Chi *et al.*, 2003), and usability guidelines developed for Web designers (Nielsen, 2003, 2004; Spool *et al.*, 2004). We will also compare SNIF-ACT to two existing models of user-WWW interactions called MESA (Miller & Remington, 2004) and CoLiDeS (Kitajima *et al.*, 2005) in the Discussion section.

1.1. Overview of the article

In the next section, we will briefly review the theories behind the SNIF-ACT model. We will focus on the underlying theories governing how the model measures information scent and consequently selects the appropriate actions based on the currently attended information content. Based on the theories, we will discuss the details of the model and the user-tracing architecture that we used to analyze the human and model data. We will then present two versions of the model. First, we will describe the details of SNIF-ACT 1.0, which was tested against a data set collected by Card *et al.* (2001) in a controlled experiment involving a small number of subjects. The purpose of that experiment was to provide detailed data on moment-to-moment user-Web interactions including keystroke data, eye-movement data, and concurrent verbal reports. This detailed set of protocols allowed us to directly test and fine-tune the basic parameters and mechanisms of SNIF-ACT 1.0. We also compared the SNIF-ACT 1.0 to a Position model that decides which link to select based solely on the position of links on a Web page. Although SNIF-ACT 1.0 provides a better fit to the data than the Position model, we also found that link selections seem to depend on the dynamic interaction between information scent and the position of the link on a Web page. We therefore extended the model to include a Bayesian satisficing mechanism that dynamically decides which link to follow and when to leave a Web page as the model sequentially evaluates link texts on a Web page. SNIF-ACT 2.0 is therefore more flexible and adaptive to the dynamic interactions between the user and different Web sites. The flexibility and adaptiveness of SNIF-ACT 2.0 make it suitable to explain aggregate user behavior across different Web Sites. Indeed, Monte Carlo simulations of the SNIF-

ACT 2.0 model showed good fits to a data set collected by Chi et al. (2003) in a controlled study involving 74 users working on tasks in realistic settings.

2. THEORY

SNIF-ACT is a model developed within Information Foraging Theory (Pirolli & Card, 1999), which employs the rational analysis method (e.g., Anderson, 1990). Pirolli's (2005) rational analyses of information foraging on the Web focused on some of the problems posed by the general task environment of Web users, and the structure and constraints of the information environment on the Web. SNIF-ACT provides a mechanistic implementation that approximates the rational analysis model. In developing the SNIF-ACT computational cognitive model, additional constraints coming from the cognitive architecture must be addressed. In particular, SNIF-ACT must employ satisficing (suffices to satisfy a particular aspiration level without maximizing, see Simon, 1955) and learning from experience. These mechanisms arise as solutions to limits on computational resources and amount of available information that are not necessarily considered constraints in rational analyses. In this section, we provide a summary of Information Foraging Theory, the rational analysis of Web foraging, and the spreading activation model of information scent that is implemented in SNIF-ACT.

2.1. Information Foraging Theory

Information foraging theory (Pirolli & Card, 1999) assumes that people develop information-seeking strategies that optimize the utility of information gained in relation to the cost of interaction. This approach shares much with the rational analysis methodology initiated by Anderson and his colleagues (Anderson, 1990; Oaksford & Chater, 1998). The rational analysis approach involves a kind of reverse engineering in which the theorist asks (a) *what* environmental problem is being solved, (b) *why* is a given behavioral strategy a good solution to the problem, and (c) *how* is that solution realized by cognitive mechanisms. The products of this approach include (a) characterizations of the relevant goals and environment, (b) mathematical rational choice models (e.g., optimization models) of idealized behavioral strategies for achieving those goals in that environment, and (c) computational cognitive models. Rational analysis is a variant form of an approach called *methodological adaptationism* that has also shaped research programs in behavioral ecology (e.g., Mayr, 1983; Stephens & Krebs, 1986; Tinbergen, 1963), anthropology (e.g., Winterhalder & Smith, 1992), and neuroscience (e.g., Glimcher, 2003).

Pirolli's (2005) rational analysis of information foraging on the Web focused on the problems of (a) the choice of the most cost-effective and useful browsing actions to take based on the relation of the navigation cues (information scent) to a user's information need and (b) the decision of whether to continue at a Web site or leave based on ongoing assessments of the site's potential usefulness

and costs. Rational choice models, and specifically approaches borrowed and modified from optimal foraging theory (Stephens & Krebs, 1986) and microeconomics (McFadden, 1974), were used to predict rational behavioral solutions to these problems. Pirolli (2005) argued that the cost-benefit assessments involved in the solution to these problems facing the Web user could be grounded in a rational utility model implemented as a spreading activation process. Activation from representations of information scent cues spreads to the user's information goal. The amount of activation received by the user's goal reflects the expected utility of choosing navigation actions associated with those cues. This spreading activation model is discussed in the next subsection.

2.1.1. *Spreading Activation and Information Scent*

SNIF-ACT employs a spreading activation mechanism to assess the utility of navigational choices. Spreading activation is assumed to operate on a large associative network that represents the Web user's linguistic knowledge. These spreading activation networks are central to SNIF-ACT, and one would prefer that they be predictive in the sense that they are (a) general over the universe of tasks and (b) not estimated from the behavioral data of the users being modeled. SNIF-ACT assumes that the spreading activation networks have computational properties that reflect the statistical properties of the linguistic environment (Anderson & Milson, 1989; Landauer & Dumais, 1997). These networks can be constructed using statistical estimates obtained from appropriately large and representative samples of the linguistic environment. Consequently, SNIF-ACT predictions for Web users with particular goals can be made using spreading activation networks that are constructed *a priori* with no free parameters to be estimated from user data.

Figure 1 presents a schematic example of the information scent assessment subtask facing a Web user. It assumes that a user has the goal of finding information about "medical treatments for cancer," and encounters a Web link labeled with the text that includes "cell", "patient", "dose", and "beam". The user's cognitive task is to predict the likelihood that a distal source of content contains desired information based on the proximal information scent cues available in the Web link labels. Pirolli (2005) presents a rational analysis (in terms of a Bayesian analysis) of the assessment problem exemplified in Fig. 1 which arrives at a spreading activation model.

----- INSERT Figure 1 ABOUT HERE -----

The spreading activation model of information scent in SNIF-ACT assumes that activation spreads from a set of cognitive structures that are the current focus of attention through *associations* to other cognitive structures in memory. Using ACT-R terminology, these cognitive structures are called *chunks* (Anderson & Lebiere, 2000). Chunks representing information scent cues are presented on the right side of Fig. 1, chunks representing the user's information need are presented

on the left side, and associations are represented by lines. The associations among chunks come from past experience. The strength of associations reflects the degree to which proximal information scent cues predict the occurrence of unobserved features. For instance, the word “medical” and “patient” co-occur quite frequently and they would have a high strength of association. Greater strength of association produces greater amounts of activation flow from one chunk to another.

Expressing the spreading activation model in the context of a user evaluating the utility of a link L on a Web page to his or her information goal G , the activation of a chunk i in the information goal is A_i , where

$$A_i = B_i + \sum_{j \in L} W_j S_{ji} . \quad (\text{Eqn 1: Activation equation})$$

In the activation equation above, B_i is the base-level activation of chunk i , S_{ji} is the association strength between chunk j representing a cue in the link L and the goal chunk i , and W_j reflects the attentional weight the model puts on chunk j . As noted in Pirolli (2005), S_{ji} is a very near approximation of what is known as *Pointwise Mutual Information* (PMI) in the information retrieval and statistical natural language literature (e.g., Manning & Schuetze, 1999). The activation equation is interpreted as a Bayesian prediction of the relevance of chunk i in the context of the chunks in the link on a Web page to which the model is currently attending (Pirolli & Card, 1999). B_i reflects the log prior odds of chunk i occurring in the world, and S_{ji} reflects the log likelihood ratio of chunk j occurring in the context of word i . The information scent of the link L is simply the sum of activations of all chunks in the information goal G

$$\begin{aligned} IS(G, L) &= \sum_{i \in G} A_i \\ &= \sum_{i \in G} (B_i + \sum_{j \in L} W_j S_{ji}) . \end{aligned} \quad (\text{Eqn 2: Information scent equation})$$

For tasks in which the information goal remains constant throughout the task—such as the tasks modeled in this paper—the base-level activations B_i can be ignored. This is because the goal chunks i remain the same throughout the task. Consequently, the base-level activations of the goal, B_i , of goal chunks do not change regardless of the link chunks j . Consequently, in the SNIF-ACT model we set B_i to zero.

The model also must deal with the case in which a link chunk j is the same as goal chunk i (e.g., if a person were looking for “medical information” and saw the word “medical” on a link). In cases of direct overlap between the information goal of the user and the information scent cues of the link (i.e., when $S_{ji} = S_{ii}$), S_{ji} reflects the log prior odds of the goal chunk i . This has the effect of making the activation equation especially sensitive to direct overlaps between information goals and information scent cues.

The model also requires the specification of the attentional weight parameter W_j . We have simply assumed that the attention paid to an individual information scent cue decays exponentially as the total number of cues increases. Specifically, we set

$$W_i = W e^{-dn}, \quad (\text{Eqn 3: Attentional weight equation})$$

where n is the number of words in the link, W is a scaling parameter, and d is a rate of decay parameter. The exponential decay function is used to ensure that the activation will not increase without bounds with the number of words in a link. Specifically, as the number, n , of words on a link gets larger, the total summed amount of attention grows to an asymptote

$$\sum_{i=1}^n W_i.$$

Exploration of the parameters suggested that we use $W = 0.1$ and $d = 0.2$ throughout the simulations. Using these parameters, we get a growth function for $\sum W_i$ that shows no substantial change (less than 1%) after $n = 20$ words (Spool *et al.*, 2004).

In order to calculate the information scent of a link on a Web page given the information goal of the user, we need to estimate S_{ji} . As discussed in Pirolli and Card (1999), it is possible to automatically construct large spreading activation networks from on-line text corpora, and calculate the estimates of S_{ji} for different words and information goals. Specifically, base-rate frequencies of all words and pairwise co-occurrence frequencies of words that occur within some distance of one another can be computed from large text corpora to estimate S_{ii} and S_{ji} . For SNIF-ACT 1.0 we obtained these estimates from a local Tipster document corpus (Harman, 1993) with a back-off to search engine queries of the Web to obtain statistics about words not contained in the Tipster collection. In SNIF-ACT 2.0 we employed estimates from locally stored samples of Web documents plus a back-off technique that queried the Web for statistics about words not present in the local Web collection (Farahat *et al.*, 2004). This general method of using a local sample of documents for most estimates plus queries to the Web as a back-off technique combines efficiency (most of the encountered words will be in the local store and statistics can be rapidly computed) with coverage (low frequency words can typically be found on the Web). Practically, PMI scores can be calculated efficiently (Farahat, Pirolli, and Markova, 2004), and theoretically, Farahat *et. al* showed that PMI scores were as least as good or better than Latent Semantic Analysis (LSA) in providing good fits to human word similarity judgments in a variety of tasks. All “stop words” such as “the” and “a” as listed in Callan, Croft, & Harding (1992) were removed from all processing.

2.2. Utility Calculations

SNIF-ACT uses spreading activation to calculate the information scent provided by words associated with links on a Web page, according to the equations specified above. These information scent values are used to evaluate the utility of actions including attending to links, selection of links, going back to a previous page within a Web site, and leaving a Web site. The specific utility calculations used in SNIF-ACT 1.0 were developed on the basis of random utility models in economics (McFadden, 1974) and stochastic models of search in optimal foraging theory (McNamara, 1982). These utility calculations were refined in SNIF-ACT 2.0 to implement satisficing (Simon, 1955, 1956). The details of these utility calculations are discussed separately below in the context of each model.

3. SNIF-ACT

A model called SNIF-ACT (Pirolli & Fu, 2003) was developed based on the theory of information scent described above (this earlier presentation of the model covered parts of SNIF-ACT 1.0). In this article we present old and new data and the newest version of the model. The basic structure of the model is shown in Figure 2. Similar to ACT-R models, SNIF-ACT has two memory components – the declarative memory component and the procedural memory component. Elements in the declarative memory component can be contemplated or reflected upon, whereas elements in the procedural memory component are tacit and directly embodied in physical or cognitive activity. Next, we will discuss each of the memory components separately and give an example showing the flow of the model as shown in Figure 2.

---- INSERT Figure 2 ABOUT HERE -----

3.1. Declarative Knowledge

Declarative knowledge corresponds to “facts about the world”, which are often verbalizable. In the current context, declarative knowledge consists of the content of Web links or the functionality of browser buttons, and the current goal of the users (e.g., evaluating a link, choosing a link, etc.). Since the current goal of SNIF-ACT is not to model how users learn to use the browser, we assume that the model has all the knowledge necessary to use the browser, such as clicking on a link, or clicking on the “back” button to go back to the previous Web page. We also assume that users have perfect knowledge of the addresses of most popular Web search engines. Declarative knowledge is pre-defined in the model in all the simulations, and does not change throughout the simulations.

3.2. Procedural Knowledge

Procedural knowledge corresponds to “how to do it” knowledge. In contrast to declarative knowledge, procedural knowledge is often not verbalizable. As in ACT-R, procedural knowledge is represented as production rules, which are represented as condition-action pairs. Table 1 shows the set of production rules in SNIF-ACT, presented in their English-equivalent forms. A production rule has a condition (IF) side and an action (THEN) side. When all the conditions on the condition side are matched, the production may be fired and when it does, the actions on the action side of the production will be executed. At any point in time, only a single production can fire. When there is more than one match, the matching productions form a “conflict set”. One production is then selected from the conflict set based on the Random Utility Model (RUM, details later), with the measure of information scent as the major variable controlling the likelihoods of selecting any one of the productions in the conflict set.

3.3. Selection of Actions

Actions of the models are represented as production rules as shown in Table 1. An example trace of the model is shown in Table 2, which shows the sequential execution of productions in the model. The model always starts with the goal of going to a particular Web site (usually a search engine) on the internet. There are two ways the model could go to a Web page, it could type the URL (Uniform Resource Locator) address, or it could use the “bookmark” pull-down menu in the browser. Since the major predictions of the model were on behavior contingent on the links displayed on a Web page, we are agnostic about the first Web sites users preferred (which are selected based on their prior knowledge rather than influenced by the information displayed on a Web page) and how they reached the Web sites of their choices to start their tasks. We therefore force the model to match users’ choices (details of this procedure are discussed in the next section). There were three major productions that competed against each other when the model was processing a Web page: Attend-to-Link, Click-Link, and Leave-Site³. Each of these productions has a utility value, which is calculated based on the measures of information scent of the links on the Web page. At any moment, the choice of these productions depended on their utility values. We will describe the calculations of the utility values with each model.

----- INSERT Table 1 and Table 2 ABOUT HERE -----

³ Since subjects stayed in the same Web site throughout the whole task in Experiment 2, the Leave-Site was only used in Experiment 1.

3.4. User-Tracing Architecture

User trace data consists of several kinds of data recorded and analyzed by our instrumentation package. Performance on the tasks was recorded using an instrumentation package that included: (a) WebLogger (Reeder, Pirolli, & Card, 2001), which is a program that tracks user keystrokes, mouse-movements, button use, and browser actions, (b) an eye tracker, and (c) video recordings that focused on the screen display. Details of the instrumentation used are given in Card, et. al (2001). WebLogger also saves the actual Web content (i.e. the text, images, scripts, etc.) that a user looked at during a browsing session. It does this by saving a cache of all pages and associated content that was viewed by the user. Eye-movements are handled by our WebEyeMapper system, which maps fixations to individual web elements (e.g., a link text) and stores the mapping in a database. Videotapes of users thinking aloud provide additional data about users' goals and subgoals, attention, and information representation (Ericsson & Simon, 1984). The video plus WebLogger and WebEyeMapper data are used to produce a Web Protocol Transcript. The Web Protocol Transcript includes interactions recorded by the WebLogger, transcribed audio/video data, and model coding of the inferred cognitive action that is associated with the data. The protocol analysis provides data that are not available from WebLogger and WebEyeMapper, especially the users' reading and evaluation of content and links.

Figure 2 shows how the User Tracer controls the SNIF-ACT simulation model and matches the simulation behavior to the user trace data (each step is indicated by a circle in Figure 2):

1. Parse the Interface Objects, Coded Protocol, and Event Log to determine the next display state and the next user action that occurs at that display state.
2. If the display state has changed, then indicate this to the SNIF-ACT system. SNIF-ACT contains production rules that actively perceive the display state and update declarative memory to contain chunks that represent the perceived portions of the display.
3. Run SNIF-ACT so that it runs spreading activation to identify the active portion of declarative memory and matches productions against working memory to select a conflict set of production rules.
4. SNIF-ACT evaluates the productions in the conflict set using the information scent computations. At the end of this step, one of the rules in the conflict set will be identified as the production to execute.
5. Compare the production just selected by SNIF-ACT to the next user action and record any statistics (notably whether or not the production and action matched). If there is a match, then execute the production selected by SNIF-ACT. If there is a mismatch, then select and execute the production that matches the user action.
6. Repeat Steps 1 - 5 until there are no more user actions.

The User-Tracing architecture was used to compare and evaluate the SNIF-ACT models. However, because there were significant differences between the two versions of SNIF-ACT, the evaluation methods were also different and are discussed in the next sections.

4. SNIF-ACT 1.0⁴

SNIF-ACT 1.0 was tested against detailed data from a small set of subjects studied in Card et al. (2001). These data allowed us to test and adjust parameters of our model to provide descriptions of user behavior. The main goal of developing SNIF-ACT 1.0 was to test the basic predictions about navigation choice behavior based on the theory of information scent discussed above. SNIF-ACT 1.0 assumes that users assess all the links on a page before making a navigation choice. To preview our results, we found that selection of links seem to be sensitive to their position on the web page. The results led us to refine our model to SNIF-ACT 2.0, in which we incorporated mechanisms from the Bayesian satisficing model (Fu & Gray, 2006; Fu, in press) that combine the measure of information scent and the position of links on the web page into a satisficing process that determines which link to select.

4.1. Tasks and Users

Tasks for the Card et al. (2001) study were modified versions of tasks compiled in a survey of 2188 Web users (Morrison, Pirolli, & Card, 2001). The two tasks analyzed in detail were:

Antz Task: After installing a state of the art entertainment center in your den and replacing the furniture and carpeting, your redecorating is almost complete. All that remains to be done is to purchase a set of movie posters to hang on the walls. Find a site where you can purchase the set of four Antz movie posters depicting the princess, the hero, the best friend, and the general.

City Task: You are the Chair of Comedic events for Louisiana State University in Baton Rouge, LA. Your computer has just crashed and you have lost several advertisements for upcoming events. You know that The Second City tour is coming to your theater in the spring, but you do not know the precise date. Find the date the comedy troupe is

⁴ Some of the results of SNIF-ACT 1.0 have been reported in Pirolli & Fu (2003), although additional description and analyses are included here.

playing on your campus. Also find a photograph of the group to put on the advertisement.

Four users were solicited from PARC and Stanford. Users were encouraged to perform both tasks as they would typically, but they were also instructed to think out loud (Ericsson & Simon, 1984) as they performed their tasks. Data from the users and tasks analyzed by Card et al. (2001) were simulated by SNIF-ACT 1.0 to produce the model fits discussed below. All stop words were removed from the description of the user tasks to calculate information scent of link texts.

Figure 3 shows examples of behavior extracted from the two tasks performed by one of the four study subjects. The behavior is plotted as a Web Behavior Graph (WBG), which is a version of a problem behavior graph (Newell and Simon, 1972). Each box in the diagram represents a state in a problem space. Each arrow depicts the execution of an operator, moving the state to a new state. Double vertical arrows indicate the return to a previous state, augmented by the experience of having explored the consequences of some possible moves. Thus time in the diagram proceeds left to right and top to bottom. Different shades surrounding the boxes in Figure 3 represent different Web sites. An X following a node indicates that the user exceeded the time limits for the task and that it was therefore a failure. The WBG in Figure 3, and the WBGs for the remaining study subjects and users, is presented in greater detail elsewhere (Card et. al, 2001). The WBG is particularly good at showing the structure of the search. One may characterize task difficulty in terms of the branchiness of the WBGs, with more branches indicating that search paths were abandoned and the user returned to a prior state. Another way of characterizing task difficulty is by the number of states visited by users. From Figure 3 it is evident that the ANTZ task is more difficult than the CITY task. This was true over all four users. The goal of SNIF-ACT 1.0 is to assess how much of the variability of the Web behavior, such as that depicted in Figure 3, is predictable from the measure of information scent.

----- INSERT Figure 3 ABOUT HERE -----

The predictions made by the SNIF-ACT 1.0 model were tested against the log files of all data sets. The model predicts two major kinds of actions: (1) which links on a Web page people will click on, and (2) when people decide to leave a site. These two actions were therefore extracted from the log files and compared to the predictions made by the model. We call the first kind of actions *link selections*, which were logged whenever a subject clicked on a link on a Web page. The second kind of actions was called *site-leaving actions*, which were logged whenever a subject left a Web site (and went to a different search engine or Web site). The two kinds of actions made up 72% (48% for link-following and 24% for site-leaving actions) of all the 189 actions extracted from the log files. The rest of the actions consisted of, for example, typing in the URL to go to a particular Web site, or

going to a pre-defined bookmark. These actions were excluded as they were more influenced by prior knowledge of the users rather than information displayed on the screen.

4.2. Utility Calculations

4.2.1. Link Selection and the Random Utility Model (RUM)

As discussed above, the spreading activation theory calculation of information scent reflects the likelihood that the link (a proximal cue) will eventually lead to the information goal (distal information). SNIF-ACT 1.0 assumes that all links on a page are sequentially processed by a user, and that production instantiations for selecting each processed link (the Click-Link production in Table 1) compete with one another. The utility of these Click-Link instantiations is calculated using the information scent equation (Eqn 2) presented above. The probability that a particular Click-Link production is selected and executed is calculated using a kind of Random Utility Model (McFadden, 1974, and see Appendix A). Consider the case in which the model is faced with a conflict set C of k Click-Link productions. The information scent for the n^{th} link is calculated by $IS(G,n)$ specified in the definition of information scent (since the goal stays the same in all our tasks, we will simply refer the information scent as $IS(n)$ from now on). Assuming that the noise parameters, ε , are independent random variables following a Gumbel distribution, the probability that link n will be chosen can be represented as a conditional probability $P(n, C)$, where

$$P(n, C) = \Pr(n | C) = \frac{e^{IS(n)/\tau}}{\sum_{j \in C} e^{IS(j)/\tau}} \quad (\text{Eqn 4: Conflict resolution equation})$$

and where $\tau = \sqrt{2} \varepsilon$ is a scaling parameter and the summation is for all j production instantiations in the conflict set C .

There are a number of points to make about the conflict resolution equation. First, as with other well-known choice equations in psychology (e.g., Luce, 1959; Thurstone, 1927) the choice of a particular link n is conditional on the utilities of other links. This means that a particular link with a particular information scent score (which determines the numerator of the conflict resolution equation) will have a probability of selection that can be high or low depending on the information scent of competing links (which determine the denominator of the same equation). Second, the size of the conflict set (the number of competing links) will affect the selection of any particular link for similar reasons. Third, as τ decreases, the model is more likely to choose the link with the highest

information scent. This is because τ is related to the variance of the noise parameter in the information scent equation. We set $\tau = 1.0$ throughout the simulations.

4.2.2. *Leaving a Patch of Information*

The SNIF-ACT 1.0 model assumes that the decision of whether or not to continue at a Web site is similar to a class of foraging decisions modeled in the optimal foraging literature (Pirolli, 2005). One of the major predictions of food foraging models concerns the time when the forager will leave a food patch. For example, in the stochastic food foraging models by McNamara (1982), a potential function $h(x)$ is defined for a given state within a food patch, x , and the optimal forager is one that maximizes the potential function. In particular, the potential function is defined as

$$h(x) = U(x) - C(t), \quad (\text{Eqn 5: Potential function equation})$$

where $U(x)$ is the utility of continued foraging in the current patch x , t is the expected amount of time that will be spent foraging in the patch, and $C(t)$ is the opportunity cost of foraging for t amount of time. McNamara's model defines the opportunity cost function as

$$C(t) = R_m * t, \quad (\text{Eqn 6: Opportunity cost function equation})$$

where R_m is the average long-term rate of gain of foraging in various food patches. The optimal policy for leaving a patch is when $h(x) < 0$, or when

$$U(x)/t < R_m. \quad (\text{Eqn 7: Patch-leaving policy})$$

SNIF-ACT 1.0 assumes that collections of Web pages form information patches.⁵ When the current utility of finding information on a Web page is perceived to be lower than the long term average utility of similar tasks on the Web, the optimal decision is to leave the Web page and pursue a different navigation path to find the information.

An implicit assumption of this optimal foraging model is that the forager has perfect knowledge of the environments (e.g., knowledge of $U(x)$, t , $C(t)$, and R_m). Similarly SNIF-ACT 1.0 makes the strong assumption that a Web user has perfect knowledge of these values. A more realistic assumption is that the forager estimates these characterizations of the Web based on experience with previous Web pages on similar Web sites. The SNIF-ACT 2.0 model discussed later incorporates a rational learning model to estimate Web properties similar to these.

⁵ For example, Eiron and McCurley (2003) show that the link structure of most Web sites will tend to form localized hierarchical structures, which is similar to the structures of food patches found in the natural environment (see Pirolli & Card, 1999).

4.3. Results

4.3.1. Link selections

The SNIF-ACT 1.0 model was matched to the link selections extracted from 8 sets of data (2 tasks X 4 subjects). The user trace comparator was used to compare each action from each subject to the action chosen by the model. Whenever a link selection was encountered, the SNIF-ACT 1.0 model ranked all links on the Web page according to the information scent of the links. We then compared the links chosen by the subjects to the predicted link rankings of the SNIF-ACT 1.0 model. If there were a purely deterministic relationship between predicted information scent and link choice, then all users would be predicted to choose the link with the smallest rank number. However, as discussed earlier, we assume that the scent-based utilities are stochastic and subject to some amount of variability due to users and context. Consequently we expect the probability of link choice to be highest for the links ranked with the greatest amount of scent-based utility, and that link choice probability is expected to decrease for links with higher rank number as determined on the basis of their scent-based utility values.

To highlight the importance of the information scent measure in the model, the ranks produced by SNIF-ACT 1.0 were compared to those produced by an alternative model that selects links based solely on their positions on the page. This model was motivated by recent findings that people tend to scan a Web page from top to bottom, and was found to be biased in selecting links at the top of a page containing Web search results (Joachims, et al., 2005). In this alternative model, the rank of a link is simply determined by its position on the web page, so that a link at the top of the page will be ranked 1, and the rank number increases as the model goes down from top to bottom of the web page. We call this model the “Position” model. Figure 4 shows the frequency distribution of the 91 link-following actions by the subjects plotted against the ranks of the links calculated by the SNIF-ACT 1.0 and the Position model. For SNIF-ACT 1.0, links that had a low rank number (i.e., high on scent-based utilities) tended to be chosen over links that had a higher rank number, indicating that link choice is strongly related to scent-based utility values. For example, Figure 4 shows that the link with the highest information scent as calculated by SNIF-ACT 1.0 was select 19 times by the subjects, and the link with the next highest score was selected 15 times by the subject. The predictive value of the model lies on the high frequencies of links on the left side of Figure 4, which slope down and level off to the right side of the figure. This result replicates a similar analysis made by Pirolli and Card (1999) concerning the ACT-IF model prediction of cluster selection in the Scatter/Gather browser, in which the rankings made by the model (which were also based on the same scent-based utilities) correlated well with the selection by the users.

For the Position model, the ranks in Figure 4 indicated the positions of the links on the Web page. Links on the top of a page will have a smaller rank number than those at the bottom; in cases where

there were more than two links on the same line, links on the left will have a lower rank number than those on the right. By this method, we found that the first link on the Web page was selected two times by the subjects, and the second link on the Web page was selected 3 times by the subjects. The frequencies of link choices increased with rank number (i.e., position on the web page) and peaked at approximately the fourth link, but after that they decreased slowly for links further down the page. The results indicated that although subjects did not simply choose the first link on a web page, there was still a higher tendency to choose links at the top of the page than those towards the bottom of the page. Indeed, for both SNIF-ACT 1.0 and the Position model, the downward trends across ranks were significant (slope = -0.32 and -0.20, $t(1,28) = 4.61$ and 6.84 respectively), suggesting that both models successfully predicted the general link-selection trends. In other words, both information scent and position on a web page have some predictive power of link selection, however, the significantly more negative slope by SNIF-ACT 1.0 indicated that the measure of information scent has more predictive power than position on a web page ($\chi^2(30) = 53.59$, $p < 0.005$). On the other hand, previous research on the predictive power of link location have focused on Web search results, and our results showed that the predictive power is still significant even in general Web pages.

---- INSERT Figure 4 ABOUT HERE ----

4.3.2. *Site-leaving actions*

To test how well information scent predicts when people will leave a site, site-leaving actions were extracted from the log files and analyzed. Site-leaving actions were defined as actions other than link-clicking that led to a different site (e.g. when the subjects used a different search engine by typing in the URL or using an existing bookmark). The results were plotted in Figure 5. It shows the mean information scent of the four Web pages the subjects visited before they left the site (i.e. Last-3, Last-2, Last-1, and Leave-Site in Figure 5). It shows that initially the mean information scent of the Web page was high, and right before the subjects left the site, the mean information scent dropped. However, given the small number of site-leaving actions that we recorded, the difference did not reach statistical significance ($t(11)=0.61$, $p=0.56$).

Figure 5 also shows the mean information scent of the Web pages right after the subjects left the site (the dotted line in Figure 5). It shows that the mean information scent on the page right after they left the site tended to be higher than the mean information scent before they left the site. This is consistent with the information foraging theory which states that people may switch to another "information patch" when the expected gain of searching in the current patch is lower than the expected gain of searching for a new information patch. In fact, from the verbal protocols, we often found utterances like "it seems that I don't have much luck with this site", or "maybe I should try

another search engine" right before subjects switch to another site. It suggests that the drop in information scent on the Web page could be the factor that triggered subjects' decision to switch to another site.

---- INSERT Figure 5 ABOUT HERE ----

4.3.3. Summary of results

We show that links chosen by the subjects were largely predicted (as reflected by the low rank numbers) by SNIF-ACT 1.0. The good match between the predictions of SNIF-ACT and the data shows the predictive power of information scent in link selections. Information scent was also shown to be sensitive to when people will decide to switch to a different Web site, although the effect is not statistically significant. When subjects left a site, the average information scent of the site tended to be decreasing. The results are consistent with the notion that as people go through a sequence of Web pages, they are building up an expectation of how likely they can find the target information on the Web sites.

The results for the Position model also show that there is a weak trend for people to select links at the top of the page over those at the bottom of the page. It is, however, likely that there is a high correlation between information scent of links and their position on a Web page. This is especially likely in situations where subjects are evaluating a list of links returned from a Web search engine, as links at the top of the returned list of links tended to be more relevant to the search terms than those further down the list. Indeed we found that this correlation was high ($r=0.64$, $t(15)=1.92$, $p<0.05$). Since SNIF-ACT 1.0 simply picks the link with the highest information scent value regardless of its position on the Web page, link selections by the model are not sensitive to the position of links. To take into account the fact that both information scent and positions influence link selection, we refine our model in SNIF-ACT 2.0 so that the model will dynamically build up an expectation on how likely the target information can be found as it processes each link on a Web page sequentially. To preview our results, we found that this dynamic mechanism provides a much better match to link selections than either the Position or the SNIF-ACT 1.0 model.

5. SNIF-ACT 2.0

Results from the test of SNIF-ACT 1.0 show that the measure of information scent provides good prediction of link selections in naturalistic user-Web interactions. We also found that the simple information of link position on a web page also seems to predict link selections. The results are consistent with the idea that the link selection process involves a dynamic evaluation process that operates on both information scent and the position or sequential order of links. In SNIF-ACT 2.0,

we hypothesize that during the link selection process, current and previous experiences with different link texts and Web sites interact dynamically and influence the final selection. The learning mechanism allows the model to adapt to the specific experiences of users as they interact with different Web pages.

SNIF-ACT 2.0 has an adaptive action evaluation and selection mechanism that dynamically chooses actions based on current and previous experiences with the link texts on the Web sites. To evaluate SNIF-ACT 2.0, we expanded our data sets to include more subjects and more tasks (Chi et al., 2003). We intend to understand how the predictions of the model can be applied to explain the dynamic user-Web interactions across different Web sites and users in realistic settings. In this section, we will first discuss the tasks in the dataset by Chi et al., followed by a description of the new learning mechanism in SNIF-ACT 2.0. We will then show the results from Monte Carlo simulations of the model and how well they matched the human data.

5.1. Tasks and Users

Chi et al. (2003) were interested in validating the predictions of an automated Web usability testing system called *Bloodhound*. Chi et al. (2003) used a remote version of a usability data collection tool based on *WebLogger* (Reeder, et. al., 2001). Subjects in the Chi et al. (2003) study downloaded this testing apparatus and went through the test at their leisure in a place of their choosing. Users were presented with specific information-seeking tasks to perform at specific Web sites. We discovered that it was difficult to infer user navigation at Web sites that relied heavily on the dynamic generation of Web pages in this dataset as we could not reproduce exactly what was on these dynamic Web pages. Consequently, we chose to simulate data from tasks performed at two Web sites in the Chi et al. (2003) data set: (1) help.yahoo.com (the help system section of Yahoo!) and (2) parcweb.parc.com (an intranet of company internal information). We will refer to these sites as “Yahoo” and “ParcWeb” respectively for the rest of the article.

---- INSERT Table 3 ABOUT HERE ----

Each of these Web sites (Yahoo and ParcWeb) had been tested with a set of eight tasks, for a total of $8 \times 2 = 16$ tasks. For each site, the eight tasks were grouped into four categories of similar types. For each task, the user was given an information goal in the form of a question. The tasks developed by Chi et al. (2003) were designed to be representative of the tasks normally performed by users of the site. The tasks are presented in Table 3.

The Yahoo and ParcWeb datasets come from a total of 74 subjects distributed approximately 30 subjects in the Yahoo dataset and 44 subjects in the ParcWeb dataset. Yahoo subjects were recruited

using Internet advertising and ParcWeb subjects were recruited from PARC employees⁶. Subjects had been asked to perform the study in the comfort of their office or anywhere else they chose. Subjects could abandon a task if they felt frustrated, and they were also told that they could stop and continue the study at a later time. The idea was to have them work on these tasks as naturally as possible. Users had been explicitly asked not to use the search feature of the site, since Chi et al. (2003) were interested in predicting navigation data. This was the preferred strategy as shown by Katz & Byrne (2003). Each subject was assigned a total of eight tasks from across different sites and each task was assigned roughly the same number of times. Whenever the user wanted to abandon a task, or if they felt they had achieved the goal, the user clicked on a button signifying the end of the task. Remote WebLogger recorded the time subjects took to handle each task, the pages they accessed, and the keystrokes they entered (if any).

Of all the user sessions collected, the data were inspected to throw out any sessions that employed the site's search engine as well as any sessions that did not go beyond the starting home page. We were not interested in sessions that involved the search engine because we wanted users to find the information using only navigation. In the end, 590 user sessions were usable (358 in Yahoo; 232 in ParcWeb). Table 4 summarizes the number of usable sessions that were collected for each task.

---- INSERT Table 4 ABOUT HERE ----

In general, we found that in both sites, there were only a few (< 10) "attractor" pages visited by most of the subjects, but there were also many pages visited by fewer than 10 subjects. In fact, a large number of Web pages were visited only once in both sites. We decided that Web pages that were visited only a few times seemed more random than systematic, and were excluded from our model simulations. In the rest of the analyses, we dropped the bottom 30% of the Web pages that were least frequently visited. As a result, Web pages that were visited fewer than three times (for all subjects) in the ParcWeb site and those visited fewer than five times in the Yahoo site were excluded for model simulations. Our assumption is that predicting pages visited most often in our sample of subjects is more important in terms of validating the SNIF-ACT model

5.2. Utility calculations

Based on the SNIF-ACT 1.0 simulations, we decided to refine the model to provide more precise predictions on the dynamic user-Web interactions. We performed Monte Carlo simulations of the model and match the results to aggregates of human data. The major extension of the model in

⁶ Since ParcWeb was quite dynamic and changed quite frequently, none of the subjects was familiar with the link structures nor knew the location of the target information before the tasks even though they were Parc employees.

SNIF-ACT 2.0 is the use of an adaptive mechanism that incrementally learns from its experiences with the links and Web pages visited. We will show how the mechanism defines stochastic decision boundaries that allow SNIF-ACT 2.0 to decide when to (1) choose a link on a Web page through a satisficing process, and (2) stop evaluating links on a web page and go back to the previous Web page.

The adaptive mechanism is based on a rational analysis of link evaluation and selection on a web page. The details of the rational analysis can be found in Appendix B. As a summary, the mechanism assumes that the probability that a link will be selected is incrementally updated through a Bayesian learning framework in which the user is gathering data from the sequential evaluation (left-right then top-down) of links on a Web page (see Fu, in press; Fu & Gray, 2006). We define the perceived closeness of the target information as a weighted sum of the IS of the links encountered on the web page (details see Appendix B). This allows us to define how utilities of productions are calculated in SNIF-ACT 2.0.

As discussed earlier, the critical productions that determine which links to follow and when to go back to the previous page were Attend-to-Link, Click-Link, and Backup-a-page. Since subjects in the Chi et al. (2003) dataset stayed in the same Web site throughout the entire session, the Leave-Site production was not used. The utilities of the critical productions are updated according to the following equations:

$$\text{Attend-to-Link: } U(n+1) = \frac{U(n) + IS(link)}{1 + N(n)}$$

$$\text{Click-Link: } U(n+1) = \frac{U(n) + IS(link)}{1 + k + N(n)}$$

$$\text{Backup-a-Page: } U(n+1) = MIS(\text{Previous Pages}) - MIS(\text{links 1 to } n) - GoBackCost$$

(Eqn 8: Utility Equations)

In the equations above, $U(n)$ represents the utility of the production at cycle n , $IS(link)$ represents the information scent of the current attended link, $N(n)$ represents the number of links attended on the Web page at cycle n , $IS(\text{Best Link})$ is the highest information scent of the links attended on the Web page, k is a scaling parameter, $MIS(page)$ is the mean information scent of the links on the Web page, and $GoBackCost$ is the cost of going back to the previous page. The values of k and $GoBackCost$ were set at $k = 5$ and $GoBackCost = 5$ in the simulations. The first two equations are derived from the rational analysis of link evaluation and selection. The last equation is based on the finding in SNIF-ACT 1.0 (see Figure 5), and is consistent with the patch-leaving policy we discussed earlier. We will illustrate this point with a hypothetical example below.

---- INSERT Figure 6 ABOUT HERE ----

Figure 6 shows a hypothetical situation in which the SNIF-ACT 2.0 model is processing a Web page. We will show how the probabilities of attending to the next link, selecting a link, and leaving the Web page will change as the model interacts with this Web page. In this hypothetical Web page, the information scent (i.e., $IS(link)$ in the utility equations above) decreases from 10 to 2 from Links 1 to 5⁷. The information scent of the links from 6 onwards stays at 2. The mean information scent of the previous pages was 10 (i.e., $MIS(Previous\ page)$), and the noise parameter τ (see the conflict resolution equation) was set to 1.0. The initial utilities of all productions were set to 0. One can see that initially, the probability of choosing Attend-to-Link is high. This is based on the assumption that when a Web page is first processed, there is a bias in learning the utility of links on the page before a decision is made. However, as more links are evaluated, the utility of the production decreases (as the denominator gets larger as $N(n)$ increases), and thus, the probability of choosing Attend-to-Link decreases. As $N(n)$ increases, the utility of Click-Link increases, and in this example, the best link evaluated so far is the first link that has information scent of 10 (i.e., $IS(Best) = 10$). The implicit assumption of the model is that since evaluation of links takes time, the more links that are evaluated, the more likely that the best link evaluated so far will be selected (otherwise the time cost may outweigh the benefits of finding a better link). As shown in Figure 6, after four links have been evaluated, the probability of choosing Click-Link is larger than that of Attend-to-Link. At this point, if Click-Link is selected, the model will choose the first (best) link and the model will continue to process the next page. However, as the selection process is stochastic (see the conflict resolution equation), Attend-to-Link may still be selected. If this is the case, as more links are evaluated (i.e., as $N(n)$ increases), the probability of choosing Attend-to-Link and Click-Link decreases. On the other hand, the probability of choosing Backup-a-Page is low initially because of the high GoBackCost. However, as the mean information scent of the links evaluated (i.e., $MIS(links\ 1\ to\ n)$) on the page decreases, the probability of choosing Backup-a-Page increases. This happens because the mean information scent of the current page is perceived to be dropping relative to the mean information scent of the previous page. In fact, after eight links are evaluated, the probability of choosing Backup-a-Page becomes higher than that of Attend-to-Link and Click-Link, and the probability of choosing Backup-a-Page keeps on increasing as more links are evaluated (as the mean information scent of the current page decreases).

As illustrated in the above example, as the model attends to each of the links on the web page, the probability of selecting Attend-to-Link decreases while that of Click-Link increases (the actual probabilities are derived from the conflict resolution equation). As a result, the utility calculations

⁷ The scent values are chosen for illustration purposes only, the actual scent values are likely to be in the range from 50 to 200.

and the set of productions implement an adaptive stopping rule for when to stop evaluating the next link, in which the stopping rule depends stochastically on the dynamic interactions between past and current experiences of the links. For example, the model is more likely to stop attending to the next link as it experiences links of diminishing scent values. Similarly, since the probability of selecting Backup-a-Page increases as the model attends to each link, the model is getting more likely to stop attending to the next link or clicking on the best link. This adaptive stopping rule is consistent with the patch-leaving policy we discussed earlier. As the information scent of the links on the current Web page drops below the mean information scent of previous pages, the model is more likely to stop processing the current Web page and abandon the current path of navigation by going back to the previous page. The utility calculations also implement a satisficing process (Simon, 1956), in which links are evaluated in sequence until one is “good enough”. This is the essence of the theory of bounded rationality coined by Simon (1955, 1956). Compared to SNIF-ACT 1.0, in which we assumed that subjects evaluate all links on a page and pick the one with the highest information scent, the satisficing process in SNIF-ACT 2.0 is a more psychologically plausible mechanism. This learning mechanism also makes the model more adaptive to specific experiences of links on a Web page, and therefore makes the model more flexible to the characteristics of different Web sites.

Finally, it is important to point out that the current mechanism does not guarantee that the “best” link will be picked. The current model is therefore consistent with the concept of bounded rationality (Simon, 1956). In other words, although the Information Foraging Theory is based on the rationality framework and the optimal foraging theory, the implementation of the model does include reasonable psychological constraints that do not always imply optimal behavior.

5.3. Results

5.3.1. *Link selections*

As the utility calculations imply, when processing a Web page, the model’s prediction of which link to select depends on both the information scent and the position of the links. To test the predictions of the SNIF-ACT 2.0 model on its selection of links, we first started SNIF-ACT 2.0 on the same pages as the subjects in all tasks. The SNIF-ACT 2.0 model was then run the same number of times as the number of subjects in each task, and the selections of links were recorded⁸. After the recordings, in case SNIF-ACT 2.0 did not pick the same Web page as subjects did, we forced the model to follow the same paths as subjects. This model-tracing process was a common method for comparing model predictions to human performance (e.g., see Anderson, Corbett, Koedinger, & Pelletier, 1995, for a review). It also allows us to directly align the model simulation results with the

⁸ We also recorded the case when the model chose to go back to the previous page. Details are presented in the next subsection.

subject data. For example, if subjects clicked on a particular Web page k time, the model would also make k selection on the same web page. Since the model faced each of the Web pages the same number of times as the subjects, ideally, the number of times the links on a particular Web page were selected by the model and subjects would be equal. For example, if there were 3 links (X, Y, Z) on a Web page and subjects clicked Link X 3 times and Link Y 1 time and did not click on Link Z, the model would be presented with the same Web page 4 times and made one link selection in each of these presentations. If the model selected Link X 1 time, Link Y 2 times, and Link Z 1 time, the correlation between the subject and the model would be $r = -0.189$.

Using the same calculations, Figure 7 shows the scatter plots of the number of times the links on all Web pages were selected by the model and subjects. As illustrated by the example earlier, if the model's predictions were perfect, all points in Figure 7 should lie on the straight line that passes through the origin with a slope of 1. Figure 7 shows that in general, the model did a good job describing the data, and the model did better in describing the data in the Yahoo tasks ($R^2=0.91$) than in the ParcWeb tasks ($R^2=0.69$). In particular, in the ParcWeb site, there were many data points lying near the x- and y-axis when the model or subjects selected the link 5 times or fewer (i.e., the area near the origin), suggesting that there were many selections made by a small number of the subjects not predicted by the model, and also many selections by small number of runs of the model (because of the noisy stochastic process) not chosen by the subjects. However, even when these data points were further excluded (those selected fewer than 5 times by both the subjects and model), we still obtained a fit of $R^2=0.64$ and $R^2=0.91$ for the ParcWeb and Yahoo tasks respectively. These results show that, in general, links frequently chosen by subjects were also chosen frequently by the model for both sites. This is important because this demonstrates the ability of SNIF-ACT 2.0 to identify the links most likely chosen by the subjects across a wide range of tasks in two very different Web sites. Theoretically, the results provided further evidence supporting the claim that the measure of information scent captures the way people evaluate mutual relevance between different link texts and information goals. From a practical point of view, we consider the ability to make predictions on which links are chosen most frequently as one of the most important criteria for evaluating a usability tool. For example, designers are able to evaluate the way information is presented on a Web site (or any information structures in general) by predicting how people are able to obtain the information they want efficiently.

---- INSERT Figure 7 ABOUT HERE ----

To highlight the predictive power of SNIF-ACT 2.0, we also compared the simulation results to those produced by the Position model and SNIF-ACT 1.0. However, since the Position model only predicts the ranks of links on a given Web page based on the position of links, we need to refine

the models so that they include a stochastic action selection mechanism to select a link. For the Position model, the Backup-a-Page production was never selected, and the probabilities of choosing the productions Attend-to-Link and Click-Link were calculated as:

$$P(\text{Attend-to-Link}) = 1 - \frac{N(n)}{\text{Number of Links on the Page}}$$

$$P(\text{Click-Link}) = \frac{N(n)}{\text{Number of Links on the Page}}$$

(Eqn 9: Probabilities of production selection in the Linear model)

where $N(n)$ is the number of links attended on the Web page at cycle n . As the model attended to each link, the Information Scent value of the link was calculated, and the model kept track of the best link encountered so far. When the Click-Link production was selected, the best link would be selected. However, unlike SNIF-ACT 2.0, the probability to click on the best link depended only on the number of links attended, and did not depend on its Information Scent value.

Figure 7 also shows the same scatter plots for SNIF-ACT 1.0 and the Position model. We see that SNIF-ACT 1.0 did a reasonable job describing the data ($R^2=0.35$ and $R^2=0.62$ for the ParcWeb and Yahoo sites respectively), showing that even without taking into account the position of links, information scent still had good predictive power on link selections. For the Position model, we obtained $R^2=0.03$ and $R^2=0.45$ for ParcWeb and Yahoo respectively. Contrary to previous findings (Joachims, et al., 2005), the Position model yielded worse fits than SNIF-ACT 1.0 and 2.0. The results showed that in general, information scent seems to be a better predictor than position information.⁹

Figure 7 shows that SNIF-ACT 1.0 and the Position model were worse at identifying many of the “attractor” pages, as shown by the data points lying on or close to the x-axis. On the other hand, both SNIF-ACT 1.0 and the Position model frequently chose links that were not chosen by the subjects, as shown by the data points lying on the y-axis. By inspecting these links, we found that links chosen frequently by subjects but not by SNIF-ACT 1.0 were all encountered early on (13 out of 16 for ParcWeb and 6 out of 6 for Yahoo); on the other hand, those links chosen by SNIF-ACT 1.0 but not by the subjects had high Information Scent values, but they were mostly at the bottom of the Web page (8 out of 12 for ParcWeb and 6 out of 7 for Yahoo). The results were consistent with the assumption of the SNIF-ACT 2.0 model: subjects tended to “satisfice” on “reasonably good” links presented earlier on the Web page rather than exhaustively finding the best links on the whole Web page.

⁹ The study by Joachims et al. only focused on lists returned from search engines, and our dataset did not allow us to separate those pages from others.

5.3.2. *Going back to the previous page*

The new utility equations allow the model to predict when it will stop evaluating links and go back to the previous Web page. Going back to the previous Web page was more likely when the utility of the Backup-a-page production became comparable or higher than that of Attend-to-Link and Click-Link productions, and consequently the Backup-a-page production was more likely to be selected by the stochastic conflict resolution equation. As shown in Figure 6, as the information scent decreases and becomes much lower than the mean information scent of previous pages, the probability of choosing the Backup-a-Page production increases. To test the model's predictions, we compared the number of times the model chose to go back on a given Web page to the number of times subjects chose to go back on the same Web page. We then performed the same regression analyses as we did when we tested SNIF-ACT 2.0 predictions on link selection. We obtained $R^2 = 0.73$ and $R^2 = 0.80$ for the ParcWeb and Yahoo sites respectively (see Figure 8). Given the large number of Web pages that we analyzed, we considered that SNIF-ACT 2.0 did a good job predicting when people would stop following a particular path and go back to the previous page. In the model, when the information scent of a page dropped below the mean information scent of previous pages, the probability of going back increased. The results provided further support for the claim that people will choose to leave a page when the information scent drops, as we found in the SNIF-ACT1.0 simulations. The results showed that the satisficing mechanism provided a good descriptive account of both link selections and when people decided to leave a Web page.

5.3.3. *Successes in finding the target pages*

In the evaluation of our model, we adopted the model-tracing approach, in which we reset our model to follow the same paths if the model selected a link different from that chosen by the subjects. This approach allows us to directly align the predictions of the model to the subjects' data. However, this raises the question that the model is not truly experiencing the exact same sequences of Web pages as the subjects, and may not truly reflect the general capabilities of the model in predicting user-Web interactions. We therefore performed simulations of the model without resetting, and compared the percentages of time the model could successfully find the target Web pages to those of subjects. The goal of the simulations was to study how well the model was able to predict the likelihood for subjects to find the target information on a given web site, and thus how well the model can be applied to usability analyses of Web sites.

We performed 500 cycles of simulations of the Position model and both versions of SNIF-ACT and obtained the percentages of successes for each model. Table 5 shows the percentages of the subjects who successfully found the target Web page, as well as percentages of times each of the models found that target Web pages. There were some "easy" tasks (ParcWeb 1a, 4a, and 4b; Yahoo

1a, 2a, 3a, 4a, 3b and 4b) where most subjects found the target Web pages, but there were a few “difficult” tasks where none of the subjects found the target Web pages (ParcWeb 2a, 3a, 2b, 3b). Table 5 shows that in general, the models were worse than subjects in successfully finding the target pages in the “easy” tasks. SNIF-ACT 2.0 was closest to subject performance among the other models in tasks in these “easy” tasks, followed by SNIF-ACT 1.0, with the Position model being the worst. However, for the “difficult” tasks, SNIF-ACT 1.0 still found many of the target Web pages, while both the Position model and SNIF-ACT 2.0 failed to find the target Web pages, thus providing a better match to subject performance. This interesting result could be explained by the fact that SNIF-ACT 1.0 selected links with the highest scent regardless of their position on the Web page, and presumably some of those correct links (with possibly the highest information scent values) were at the bottom of the Web pages that both the Position model and SNIF-ACT 2.0 could not find. The good fits of SNIF-ACT 2.0 again demonstrate that the satisficing mechanism provides a good psychologically plausible account of the process of sequential evaluation of links. The results also demonstrate the general capabilities of the model to be utilized as a tool to predict task difficulties and for general usability analyses of Web sites. Usability analysts could first identify a range of typical information goals for particular Web sites or large information structures. The model can then be applied to search for these information goals using the Web site, and the percentages of successes could provide a good index of how likely users are able to find the target information in general. The good match of the model to human behavior demonstrates the validities of applying the model to conduct this kind of automatic usability analyses system. In the Discussion section, we will briefly describe such a system called Bloodhound.

5.3.4. *Summary of results*

We conclude that SNIF-ACT 2.0 did a good job predicting user-Web interactions in a wide range of users and tasks in realistic settings. In both versions of the model, SNIF-ACT 1.0 and SNIF-ACT 2.0, we found that the measure of information scent provides good descriptions of how people evaluate mutual relevance of link texts and their information goals. We also compared the models to a simple Position model that selects links based solely on their positions on the Web page. Consistent with previous results (Joachims et al., 2005), we found that the Position model did have some predictive power in characterizing link selections. On the other hand, both versions of SNIF-ACT provide much better fits to human data than the Position model, demonstrating that the measure of information scent does a much better job in predicting user-Web interactions.

To combine the predictive power of position of links and information scent, we developed SNIF-ACT 2.0, which implements a stochastic, adaptive evaluation and selection mechanism when evaluating and selecting links on a Web page. The major theoretical premise of SNIF-ACT 2.0 is derived from the assumption that, since evaluation of links takes time, the time cost incurred from

evaluating all links on a page may not be justified, and thus as links are evaluated sequentially, the selection of links will be affected by a dynamic tradeoff of the perceived likelihood of finding the target information as the model continues to evaluate the list of links and the cost incurred in doing so. Unlike SNIF-ACT 1.0, which selects the best links on a Web page regardless of its position, SNIF-ACT 2.0 satisfices on a good enough link without exhausting all links on a Web page. Our results show that SNIF-ACT 2.0 provides a better descriptive account of user-Web interactions than both SNIF-ACT 1.0 and the Position model. By developing our model on the basis of a general theoretical framework of rational analyses, our goal is to show how a more general methodology can be useful for developing a solid theoretical foundation for usability studies for a wide range of situations.

Besides link selection, SNIF-ACT 2.0 also provides good descriptions of when people will go back to the previous page. Based on results from the SNIF-ACT 1.0 simulations, the probability that the model will go back to the previous page increases as the information scent of the current page is low compared to the mean information scent of previous pages. This mechanism is based on the assumption that when the model processes a page, it develops an expectation of the level of information scent of future pages. When the information scent of a page drops below the expected level, the model is more likely to go back to the previous page. The dynamic selection mechanism therefore successfully provides an integrated account of both link selection and when people decide not to continue further on a given Web page. Indeed, when we allow the model to freely search on the Web sites, we found that SNIF-ACT 2.0 provides the best match to human data in finding (whether successful or not) the target information. This is important as it demonstrates the model's capability to predict task difficulties and how it can be extended to an automatic usability analyses tool, which we will describe in the discussion section next.

6. GENERAL DISCUSSION

Pirolli and Card (1999) presented the theory of information foraging that casts the general problem of finding information in terms of an adaptation process between people and their information environments. In this article, we extended the theory and presented a computational model that combines the Bayesian satisficing mechanism (Fu & Gray, 2006) with the random utility theory to explain user-Web interactions. In particular, we showed that the model made good predictions about link selections on a Web page and when people would abandon the current page and go back to the previous page. In two experiments, we show that the predictions match human data well at both the individual and the aggregate level. Although the model is tested only on interactions between

humans and the WWW, we believe that the fundamental principles behind the model are general enough to be applicable to other large information structures.

One of the assumptions of conventional optimal foraging models (Stephens & Krebs, 1986) is that the forager has perfect knowledge of the environment. This assumption is similar to the economic assumption of the “rational person”, who has perfect knowledge and unlimited computational resources to derive the optimal decision (Simon, 1955; 1956). Simon argued that human decision makers are better characterized as exhibiting bounded rationality – limited knowledge and various psychological constraints often make the choice process far from optimal. Instead of searching for the optimal choice, choices are often made once they are good enough based on some estimation of the characteristics of the environment – a process called satisficing. In our model, the satisficing process is implemented through competition between productions. Instead of processing all links on a page and selecting the best link, utilities of productions are updated as each link is evaluated, and once a link is found to be good enough, the model will choose it. The same mechanism is used to implement the patch-leaving policy, in which the average utility of staying on the same Web page is constantly updated and compared to the expected value estimated from previous experiences. We show that the model based on the bounded rationality framework provides good description of user-Web interactions.

As we proceeded from modeling individual to aggregate behavior, we were making predictions about the emergent behavior of the population of Web users. This approach is similar to the analyses of Web user behavior by Huberman et. al. (1997). Huberman et. al show that the distribution of the length of sequences of Web page visits can be characterized by the Inverse Gaussian distribution – a finding that they called the *Law of Surfing*. The Law of Surfing assumes that Web page visits can be modeled as a random walk process in which the expected utility of continuing to the next page is stochastically related to the expected utility of the current page. An individual will continue to surf until the expected cost of continuing is perceived to be larger than the discounted expected value of the information to be found in the future. Our model shares the same basic assumptions as those behind the derivation of the Law of Surfing, and in Appendix C, we show that the predictions of our model on aggregate behavior are consistent with those of the Law of Surfing. On the other hand, instead of predicting how many links a user will click through on the same Web site, our model is able to produce more fine-grained predictions that focus on how evaluation of content on a Web page will affect link selections and when one will go back to the previous page.

There have been other successful models for user-Web interactions, although each of them has a slightly different focus from SNIF-ACT. For example, CoLiDes (Kitajima, Blackmon, & Polson, 2000), was implemented in the Construction-Integration architecture that explains user-Web behavior on a single Web page. Another model, called MESA by Miller & Remington (2004) makes good predictions on user behavior in different tree-like Web site architectures. Each of these models

has its strength that provides strong motivation for future improvement of the SNIF-ACT model. We will provide a review of existing models of user-Web interactions in Section 6.3. In the next two sections, we will discuss the applications, limitations and future directions of the SNIF-ACT model.

6.1. Applications of the SNIF-ACT Model

From a practical point of view, computational models of user-Web interactions are expected to improve current human-information technology designs. Existing guidelines for designs often rely on a set of vague “cognitive principles” that often only provide coarse predictions about user behavior. The major advantage of using computational models is that they allow simulations of the *integration* of various cognitive processes and how they interact to affect behavior. These predictions cannot be obtained by simply applying superficial applications of vague “cognitive principles”. Another obvious advantage is that it has the potential to perform fully automatic evaluations of information structures. Given the demands in private industry and public institutions to improve the Web and the scarcity of relevant psychological theory, there is likely to be continuing demand for scientific inquiries that may improve commerce and public welfare.

One of the ongoing projects that instantiates the practical capabilities of the SNIF-ACT model is a system called Bloodhound¹⁰ (Chi et al, 2003). A person (the Web site analyst) interested in doing a usability analysis of a Web site must indicate the Web site to be analyzed, and provide a candidate user information goal representing a task that users are expected to be performing at the site. The Bloodhound system starts with a Web-crawler program that develops a representation of the linkage topology (the page-to-page links) and downloads the Web pages (content). From these data, Bloodhound analyzes the Web pages to determine the information scent cues associated with every link on every page. At this point Bloodhound essentially has a representation of every page-to-page link, and the information scent cues associated with that link. From this, Bloodhound develops a graph representation in which the nodes are the Web site pages, the vertices are the page-to-page links at the site, and weights on the vertices represent the probability of a user choosing a particular vertex given the user’s information goal and the information scent cues associated with the link. This graph is represented as a page-by-page matrix in which the rows represent individual unique pages at the site, the columns also represent Web site pages, and the matrix cells contain the navigation choice probabilities that predict the probability, based on the measure of information scent and the conflict resolution equation, that a user with the given information goal, at a given page, will choose to go to a linked page. Using matrix computations, this matrix is used to simulate user flow at the Web site by assuming that the user starts at some given Web page and iteratively chooses to go to

¹⁰ The Bloodhound system does not include a satisficing mechanism so it is similar to SNIF-ACT 1.0, but it has a better interface for users to interact with the system.

new pages based on the predicted navigation choice probabilities. The user flow simulation yields predictions concerning the pattern of visits to Web pages, and the proportion of users that will arrive at target Web pages that contain the information relevant to their tasks. As part of the Bloodhound project, an input screen is created so that Web site analysts can enter specifications of user tasks, the Web site URL, and the target pages that contain the information relevant to those tasks. An analysis is then done by Bloodhound and a report is then automatically generated that shows such measures as the predicted number of users who will be able to find target information relevant to the specified task, as well as intermediate navigation pages that are predicted to be highly visited that may be a cause of bottlenecks. Unlike the model-tracing method we used when evaluating SNIF-ACT 2.0, the system demonstrates the general capability of the model to travel to all pages on the Web site and generate a probability profile for the whole site. The development of an automatic tool that accurately models user-Web behavior will greatly facilitate the interactive process of developing and evaluating Web sites.

6.2. Cognitive Models of Web Navigation

There have been many attempts to understand Web users and to develop Web usability methods. Empirical studies (Choo *et al.*, 2000) have reported general patterns of information seeking behavior, but have not provided much in the way of detailed analysis. Web usability methodologists (Brinck *et al.*, 2001; Krug, 2000; Nielsen, 2000; Spool *et al.*, 1999) have drawn on a mix of case studies and empirical research to extract best design practices for use during development as well as evaluation methods for identifying usability problems (Garzotto *et al.*, 1998). For instance, principles regarding the ratio of content to navigation structure on Web pages (Nielsen, 2000), the use of information scent to improve Web site navigation (User Interface Engineering, 1999), reduction of cognitive overhead (Krug, 2000), writing style and graphic design (Brinck *et al.*, 2001), and much more, can be found in the literature. Unfortunately, these principles are not universally agreed upon and have not been rigorously tested. For instance, there is a debate about the importance of download time as a usability factor (Nielsen, 2000; User Interface Engineering, 1999). Such methods can identify requirements and problems with specific designs, and may even lead to some moderately general design practices, but they are not aimed at the sort of deeper scientific understanding that may lead to large improvements in Web interface design.

The development of theory in this area can greatly accelerate progress and meet the demands of changes in the way we interact with the Web (Newell & Card, 1985). Greater theoretical understanding and the ability to predict the effects of alternative designs could bring greater coherence to the usability literature, and provide more rapid evolution of better designs. In practical terms, a designer armed with such theory could explore and explain the effects of different design

decisions on Web designs before the heavy investment of resources for implementation and testing. Theory and scientific models themselves may not be of direct use to engineers and designers, but they form a solid and fruitful foundation for design models and engineering models (Card *et al.*, 1983; Paternò *et al.*, 2000). Unfortunately, cognitive engineering models that had been developed to deal with the analysis of expert performance on well-defined tasks involving application programs (e.g., Pirolli, 1999) have had limited applicability to understanding foraging through content-rich hypermedia, and consequently new theories are needed .

The SNIF-ACT model presented in this paper is one of several recently developed cognitive models aimed at a better understanding of Web navigation. Web navigation, or browsing, typically involves some mix of scanning and reading Web pages, using search engines, assessing and selecting links on Web pages to go to other Web pages, and using various backtracking mechanisms (e.g. history lists or back buttons on a browser). None of these recently developed cognitive models (including SNIF-ACT 1.0) offers a complete account of all of these behaviors that are involved in a typical information foraging task on the Web. The development of SNIF-ACT has been driven by a process of rational analysis (Anderson, 1990) of the tasks facing the Web user and successive refinement of models in a cognitive architecture that is aimed to provide an integrated theory of cognition (Anderson & Lebiere, 1998). SNIF-ACT has focused on modeling how users make navigation choices when browsing over many pages until they either give up or find what they are seeking. These navigation choices involve which links to follow, or when to give up on a particular path and go to a previous page, another Web site, or a search engine. SNIF-ACT may be compared to two other recent models of Web navigation, MESA (Miller & Remington, 2004) and CoLiDeS (Kitajima *et al.*, 2005), which are summarized in the next subsections.

6.2.1. MESA

MESA (Miller & Remington, 2004) simulates the flow of users through tree structures of linked Web pages. MESA is intended to be a cognitive engineering model for calculating the time cost of navigation through alternative Web structures for given tasks. The focus of MESA is on link navigation, which empirical studies (Katz & Byrne, 2003) suggest is the dominant strategy for foraging for information on the Web. MESA was formulated based on several principles: (a) the *rationality principle*, which heuristically assumes that users adopt rational behavior solutions to the problems posed by their environments (within the bounds of their limitations), (b) the *limited capacity principle* which constrains the model to perform operations that are cognitively and physically feasible for the human user, and (c) the *simplicity principle*, which favors good approximations when added complexity makes the model less usable with little improvement in fit (see also, Newell & Card, 1985).

MESA scans the links on a Web page in serial order. MESA navigates with three basic operators that (1) assess the relevance of a link on a Web page, (2) select a link, and (3) backtrack to a previous page. MESA employs a *threshold strategy* for selecting links and an *opportunistic strategy* for temporarily delaying return to a previous page. MESA scans links on a Web page in serial order. If a link exceeds an internal threshold, it selects that link and goes to the linked page. Otherwise, if the link is below threshold, MESA continues scanning and assessing links. If MESA reaches the end of a Web page without selecting a link, it re-scans the page with a lower threshold, unless the threshold has already been lowered, or if marginally relevant links were encountered on the first scan.

MESA achieves correlations of $r^2 = 0.79$ with human user navigation times across a variety of tasks, Web structures, and quality of information scent (Miller & Remington, 2004). MESA does not, however, directly interact with the Web, which requires the modeler to hand-code the structure of Web that is of concern to the simulation. MESA also does not have an automated way of computing link relevance (the information scent of links), requiring that modelers separately obtain ratings of stated preferences for links. Both of these concerns are addressed by the SNIF-ACT model.

6.2.2. CoLiDeS

CoLiDeS (Kitajima *et al.*, 2005) is model of Web navigation that derives from Kintsch's (1998) construction-integration cognitive architecture. The CoLiDeS cognitive model is the basis for a cognitive engineering approach called CWW (Cognitive Walkthrough for the Web, Blackmon *et al.*, 2002). Construction-integration is generally a process by which meaningful representations of internal and external entities such as texts, display objects, and object-action connections are constructed and elaborated with material retrieved from memory, then a spreading activation constraint satisfaction process integrates the relevant information and eliminates the irrelevant. CoLiDeS includes meaningful knowledge for comprehending task instructions, formulating goals, parsing the layout of Web pages, comprehending link labels, and performing navigation actions. In CoLiDeS these spreading activation networks include representations of goals and subgoals, screen elements, and propositional knowledge, including object-action pairs. These items are represented as nodes in a network interconnected by links weighted by strength values. Activation is spread through the network in proportion to the strength of connections. The connection strengths between representations of a user's goal and screen objects correspond to the notion of information scent. As discussed below, these strengths are partly determined by Latent Semantic Analysis measures (LSA, Landauer & Dumais, 1997).

Given a task goal, CoLiDeS (Kitajima *et al.*, 2005) forms a *content subgoal* representing the meaning of the desired content, and a *navigation subgoal* representing the desired method for

finding that content (e.g., “use the Web site navigation bar”). CoLiDeS then proceeds through two construction-integration phases: (1) an *attention phase*, which determines which display items to attend to, and (2) an *action-selection phase* which results in the next navigation action to select. During the attention phase, a given Web page is parsed into subregions based on knowledge of Web and GUI layouts, knowledge is retrieved to elaborate interpretations of these subregions, and constraint satisfaction selects an action determining the direction of attention to a Web page subregion. During the action selection phase, representations of the elements of the selected subregion are elaborated by knowledge from long-term memory. The spreading activation constraint satisfaction process then selects a few objects in the subregion as relevant. Another constraint satisfaction process then selects eligible object-action pairs that are associated with the relevant items. This determines the next navigation action to perform.

In both the attention phase and the action-selection phase, spreading activation networks are constructed, activation is spread through the networks, and the most active elements in the network are selected and acted upon. As noted above, LSA is used to determine the relevance (information scent) of display objects to a user’s goal. LSA is a technique, similar to factor analysis (principal components analysis), computed over a word by document matrix tabulating the occurrence of terms (words) in documents in a collection of documents. Terms (words) can be represented as vectors in a factor space in which the cosine of the angle between those vectors represents term-to-term similarity (Manning & Schuetze, 1999), and those similarity scores correlate well with such things as judgments of synonymy (Landauer, 1986). In CoLiDeS, relevance is determined by five factors (Kitajima et al., 2005) : (1) semantic similarity as measured as the cosine of LSA term vectors representing a user’s goal and words on a Web page, (2) the LSA term vector length of words on a Web page, which is assumed to measure the familiarity of the term, (3) the frequency of occurrence of terms in document collection on which LSA has been computed, (4) the frequency of encounter with Web page terms in a user’s session, and (5) literal matches between terms representing the user’s goal and the terms on a Web page. These five factors combine to determine the strengths of association among elements representing goal elements and Web page elements, which determines the spread of activation and ultimately the control of attention and action in CoLiDeS.

The primary evaluation of CoLiDeS comes from an Web usability engineering model called Cognitive Walkthrough for the Web (CWW, Blackmon et al., 2002; Kitajima et al., 2005). CWW is used to find and identify usability problems on given Web pages. This includes prediction of the total number of clicks to accomplish a goal (a measure of task difficulty), and the identification of problems due to lack of familiar wording on Web pages, links that compete for attention, and links that have weak information scent.

6.2.3. *Relations between SNIF-ACT and other models*

SNIF-ACT, like MESA, is a simulation of how users navigate over a series of Web pages, although SNIF-ACT is not artificially restricted to tree-like structures and deals with actual Web content and structures. Similar to MESA, SNIF-ACT is founded on a rational analysis of Web navigation, although the rational analysis of SNIF-ACT derives from Information Foraging Theory (Pirolli, 2005; Pirolli & Card, 1999). This rational analysis guides the implementation of SNIF-ACT as a computational cognitive model. The initial implementations of SNIF-ACT have implicitly assumed a slightly different version of MESA's simplicity principle: SNIF-ACT was developed under the assumption that the complexity of Web navigation behavior could best be addressed by a process of successive approximation. This involves first modeling factors that are assumed to control the more significant aspects of the behavioral phenomena and then proceeding to refine the model to address additional details of user behavior.

As argued elsewhere (Pirolli, 2005), the use of information scent to make navigation choices during link following on the Web is perhaps the most significant factor in determining performance times in seeking information. This is because navigation through a Web structure, such as a Web site, can be characterized as a search process over a graph in which graph nodes represent pages and graph edges represent links among pages. Although the underlying structure is a graph, the observed search process typically forms a tree. Each search tree node, representing a visited page, has some number of branches emanating from it, corresponding to the links emanating from that page to linked pages. If the user makes perfect navigation choices at each node, only one branch is followed from each node in the tree along the shortest path from a start node (representing a starting Web page) to a target node (representing a page satisfying the user's goal). Performance times will be proportional to the length of that minimal path. On the other hand, if the quality of information scent does not support perfect navigation choices, then more than one branch will be explored from each node visited, on average. Consequently, performance times will grow exponentially with the minimum distance between the start page and the target, and the size of the exponent will grow with the average number of incorrect links followed per node (Pirolli, 2005). In the general case, small changes in information scent can cause a qualitative change from costs that grow linearly with the the minimum distance from start to target, to costs that grow exponentially with minimum distance—what has been called a *phase transition* (Hogg & Huberman, 1987) in search costs. Consequently, the development of SNIF-ACT has focused first on modeling the role of information scent in navigation choice. In this respect it is much like CoLiDeS (Kitajima et al., 2005).

However, SNIF-ACT differs in several respects from CoLiDeS. The model of information scent is based on a rational analysis of navigation choice behavior (Pirolli, 2005). The rational analysis is specified as a Random Utility Model (McFadden, 1974) that includes a Bayesian assessment of the

likelihood of achieving an information goal given the available information scent cues. Also unlike CoLiDeS, SNIF-ACT derives from the ACT-R architecture (Anderson & Lebiere, 1998). Although we currently do not make use of the full set of modeling capabilities in ACT-R, we expect those capabilities to be useful in successive refinements of SNIF-ACT. For instance, SNIF-ACT does not currently make use of ACT-R modules for the prediction of eye movements and other perceptual-motor behavior, which would be crucial to the prediction of how users scan individual Web pages and why users often fail to find information displayed on a Web page (but see Brumby & Howes, 2004). SNIF-ACT also does not make use of ACT-R's capacity for representing information-seeking plans that are characteristic of expert Web users (Bhavnani, 2002). Our choice of ACT-R as the basis for the SNIF-ACT model is partly driven by the expectation that other developed aspects of ACT-R can be used in more detailed elaborations of the basic SNIF-ACT model.

Although SNIF-ACT could not predict which Web site people would go to when they first start to search for information (by actions other than link-clicking), the model seemed to match well with human data on when they decided to go back to the previous pages. Being able to predict how long users will spend at a Web site, or on a Web foraging session, has been addressed by stochastic models of aggregate user behavior (Baldi *et al.*, 2003; Huberman *et al.*, 1998). We build upon optimal foraging models (Charnov, 1976; McNamara, 1982) to develop a rational analysis of information patch leaving (Pirolli & Card, 1999) that specifies the decision rule for abandoning the current link-following path. This rational analysis is also implemented in SNIF-ACT. To conclude, we found that while different cognitive models address slightly different aspects of user-Web interactions, there is no theoretical reason why they could not be integrated to complement each other in their strengths and weaknesses. In fact, we find the successes of these cognitive models of user-Web interactions demonstrate the promising aspect of developing a strong theoretical foundation for characterizing and understanding complex human-technology interactions.

6.3. Limitations and Future Directions

6.3.1. *Sequential vs hierarchical processing of Web pages*

One of the assumptions of the SNIF-ACT model is the sequential processing of links on a Web page. This assumption is realistic for the tasks that we analyzed, in which subjects often used search engines that returned a list of links for them to process. While we believe that this is one of the dominant modes of user-Web interactions for general information-seeking tasks, the assumption of sequential processing of links may not apply as well in certain kinds of Web pages. For example, Blackmon, Kitajima, and Polson (2005) studied how people processed Web pages that were categorized under different headings and sub-regions. They found that people tended to scan headings to identify the sub-regions of the Web page that were semantically most similar to their

user goals. Interestingly, they found that when there was a high-scent heading on the Web page, people tended to focus on the sub-region categorized under the high-scent heading and ignored the rest of the Web page. Blackmon et al.'s results implied a hierarchical, instead of sequential, processing of links on a Web page in these kinds of Web pages.

At this point, SNIF-ACT was developed at a level of abstraction that was not sensitive to different visual layouts of the Web pages, and thus could not predict results from Blackmon et al. On the other hand, the sequential processing of links in SNIF-ACT is at the evaluation stage, not at the attentional stage. Our plan is that once we have a better understanding of the relationship between people's attention process to different links and different visual layouts, it is possible to re-order the sequence of links evaluated by SNIF-ACT based on the relationship. In fact, by recording detailed eye-movements of users while they are navigating on the Web, models have been constructed that predict sequences of fixations are constructed to explain low-level perceptual processes in information seeking (Brumby & Howes, 2004; Hornof, 2004). As complex Web pages are becoming more common, a good theory of attention allocation as a function of different visual layouts is definitely important in predicting navigational behavior. Our goal is to incorporate existing results and perform further studies to understand attention allocation strategies in complex Web pages, and combine these results in future versions of the SNIF-ACT model. In fact, we believe that such a synergy will result in a more detailed and predictive model of Web navigation.

6.3.2. *Users with different background knowledge*

In both SNIF-ACT 1.0 and 2.0, we tested subjects on general information-seeking tasks that involve little domain-specific knowledge. Indeed, our model is based on weak problem-solving methods that do not depend on domain-specific knowledge. It is possible that in specific domains, for example, for Web sites that contain medical information for practitioners, expert users (either expert in the domain or in the Web sites) may perform differently by forming complicated goal structures (e.g., see Bhavnani, 2002) that possibly cannot be handled by the current version of SNIF-ACT (although it is almost trivial to implement goal structures in a production system, see Anderson & Lebiere, 1998). We do not know exactly how expertise will influence the user-Web interactions and whether the influence will have large variability across domains. The question is clearly subject to future research.

A related question is how background knowledge will affect the computations of information scent. For example, familiarities of different words for a college-level and a 9th-grade user could be very different (as they could be different between professional anthropologists and astro-physicists), and thus may affect the measurement of relatedness of two sets of words for different groups of users with very different background knowledge. One approach is to divide the text corpus a priori into sets that correspond to different groups of users with different background knowledge, and

perform the information scent calculations using these separate text corpora (e.g., see Kitajima, Blackmon, & Polson, 2005). This will allow the model to be sensitive to individual differences in background knowledge.

Another related question is how well are usability analysts able to generate typical information goals as required by the current model. The current evaluation of SNIF-ACT does assume that a well-defined information goal is presented to the user. One could imagine that in many cases, users do not have a well-formulated information goal, but rather a vague or ill-defined information goal that motivates them to search on the WWW to either understand a topic better, to acquire some conceptual framework in a particular domain, or to investigate the opinions of others on a particular topic or problems. Obviously, our model was not able to answer these questions directly, and more research is needed to understand how these information goals would arise as people are engaged in this kind of ill-defined, “sense-making” tasks.

7. ACKNOWLEDGEMENT

We thank Susan Dumais, Anthony Hornof, Marilyn Blackmon, Sarah Miller, Jennifer Tsai, and three anonymous reviewers for their comments on the manuscript. Portions of this research have been supported by an Office of Naval Research Contract No. N00014-96-C-0097 to P. Pirolli and S.K. Card and Advanced Research and Development Activity, Novel Intelligence from Massive Data Program Contract No. MDA904-03-C-0404 to S.K. Card and Peter Pirolli.

8. REFERENCES

- Ainslie, G., & Haslam, N. (1992). Hyperbolic discounting. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 57–92). New York: Russell Sage Foundation.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., Bothell, D., Byrne, M., Douglass, D., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological review*, 111 (4), 1036-1060.
- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). The cognitive tutors: lessons learned. *Journal of the Learning Sciences*, 4(2), 167–207.
- Anderson, J.R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Bhavnani, S.K. (2002). Domain-Specific Search Strategies for the Effective Retrieval of Healthcare and Shopping Information. *Proceedings of CHI'02*, 610-611.
- Blackmon, M. H., Kitajima, M., & Polson, P. G. (2005). Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. *Chi Letters*, 7: *Proceedings of CHI 2005* (ACM Press), 31-41.
- Brumby, D.P., & Howes, A. (2004). Good enough but I'll just check: web-page search as attentional refocusing. 6th Internal Conference on Cognitive Modeling, Pittsburgh, PA
- Callan, P., Croft, W. B., & Harding, S. M.. The INQUERY retrieval system. In *Proceedings of DEXA-92*, 3rd International Conference on Database and Expert Systems Applications, pages 78–83, 1992.
- Card, S., Pirolli, P., Van Der Wege, M., Morrison, J., Reeder, R., Schraedley, P., & Boshart, J. (2001). Information scent as a driver of Web Behavior Graphs: Results of a protocol analysis

method for web usability. *CHI 2001, ACM Conference on Human Factors in Computing Systems, CHI Letters*, 3(1), 498-505.

Charnov, E. L. (1976). Optimal foraging: The marginal value theorem. *Theoretical Population Biology*, 9, 129-136.

Chi, E. H., Rosien, A., Suppattanasiri, G., Williams, A., Royer, C., Chow, C., Robles, E., Dalal, B., Chen, J., & Cousins, S. (2003). The Bloodhound Project: Automating discovery of Web usability issues using the InfoScent simulator. *CHI 2003, ACM Conference on Human Factors in Computing Systems, CHI Letters*, 5(1), 505-512.

Choo, C. W., Detlor, B., & Turnbull, D. (2000). Web work: Information seeking and knowledge work on the world wide web. Dordrecht: Kluwer Academic Publishers.

Dumais, S. (2003). Data-driven approaches to information access. *Cognitive Science*, 27, 491-524.

Eiron, N., & McCurley, K. S. (2003). *Locality, hierarchy, and bidirectionality in the Web*. Paper presented at the WAW.

Ericsson, K.A. and H.A. Simon, *Protocol Analysis: Verbal reports as data*. 1984, Cambridge, MA: MIT Press.

Farahat, A., Pirolli, P., & Markova, P. (2004). *Incremental methods for computing word pair similarity* (TR-04-6-2004). Palo Alto, CA: PARC, Inc.

Fu, W.-T. (in press), Adaptive Tradeoffs between Exploration and Exploitation: A Rational-ecological Approach. Exploitation: A Rational-Ecological Approach. In Gray, W.D. (Ed), *Integrated Models of Cognitive Systems*, Oxford: Oxford University Press.. In Gray, W.D. (Ed), *Integrated Models of Cognitive Systems*, Oxford: Oxford University Press.

Fu, W.-T., & Gray, W. D. (2006). Suboptimal Tradeoffs in Information-Seeking. *Cognitive Psychology* 52, 195-242.

Garzotto, F., Matera, M., & Paolini, P. (1998). Model-based heuristic evaluation of hypermedia usability. Paper presented at the Working Conference on Advanced Visual Interfaces, L'Aquila, Italy.

- Harman, D. (1993). *Overview of the first text retrieval conference*. Paper presented at the 16th Annual International ACM/SIGIR Conference, Pittsburgh, PA.
- Hornof, A. J., & Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays. *Proceedings of ACM CHI 2003: Conference on Human Factors in Computing Systems*, New York: ACM, 249-256.
- Howes, A., Payne, S.J. & Richardson, J. (2002). An instance-based model of the effect of previous choices on the control of interactive search. In W.D.Gray and C.D.Schunn (Eds.) *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*. Fairfax, Virginia, 2002. 476-481.
- Huberman, B. A., Pirolli, P., Pitkow, J., & Lukose, R. J. (1998). Strong regularities in World Wide Web surfing. *Science*, 280, 95-97.
- Joachims, T., Granka, L., Pang, B., Hembrooke, H., and Gay, G. (2005) Accurately Interpreting Clickthrough Data as Implicit Feedback. *Proceedings of the Conference on Research and Development in Information Retrieval*, pp. 1-15.
- Katz, M. A., & Byrne, M. D. (2003). Effects of scent and breadth on use of site-specific search on e-commerce Web sites. *ACM Transactions on Computer-Human Interaction*, 10(3), 198-220.
- Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000). A comprehension-based model of Web navigation and its application to Web usability analysis. In S. McDonald, Y. Waern & G. Cockton (eds.), *People and Computers XIV—Usability or Else!*
- Kitajima, M., Blackmon, M.H. & Polson, P.G. (2005). Cognitive Architecture for Website Design and Usability Evaluation:Comprehension and Information Scent in Performing by Exploration. *HCI International 2005*.
- Krug, S. (2000). *Don't make me think*. New York: Circle.com.
- Landauer, Thomas K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science*, 10, 477-493.

- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Loewenstein, G. & Prelec, D. (1991). Negative time preference. *The American Economic Review*, 81 (2), 347-352.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers of econometrics*. New York: Academic Press.
- Miller, C. S., & Remington, R. W. (2004). Modeling information navigation: Implications for information architecture. *Human Computer Interaction*, 19(3), 225-271.
- McNamara, J. (1982). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, 21, 269-288.
- Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological review*, 108 (1), 96-112.
- Morrison, J.B., P. Pirolli, and S.K. Card. A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. In *Conference on Human Factors in Computing Systems, CHI '01*. 2001. Seattle, WA: ACM Press.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Card, S. K. (1985). The prospects for a psychological science in human-computer interactions. *Human-Computer Interaction*, 2, 251-267.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Nielsen, J. (2000). *Designing web usability*. Indianapolis, IN: New Riders.
- Paternò, F., Sabbatino, V., & Santoro, C. (2000). Using information in task models to support design of interactive safety-critical applications. Paper presented at the Working Conference on Advanced Visual Interfaces, AVI 2000, Palermo, Italy.
- Pirolli, P. & Card, S.K. (1999). Information foraging. *Psychological Review*, 106, 643-675.

- Pirolli, P. (1999). Cognitive engineering models and cognitive architectures in human-computer interaction, in F.T. Durso, et al., (Eds.) *Handbook of applied cognition*, John Wiley & Sons: West Sussex, England. p. 441-477.
- Pirolli, P. (2005). Rational analyses of information foraging on the Web. *Cognitive Science*, 29(3), 343-373.
- Pirolli, P., Fu, W., Reeder, R., & Card, S. K. (2002). A user-tracing architecture for modeling interaction with the World Wide Web. In M. D. Marsico & S. Levialdi & L. Tarantino (Eds.), *Proceedings of the Conference on Advanced Visual Interfaces*, AVI 2002 (pp. 75-83). Trento, Italy: ACM Press.
- Pirolli, P., & Fu, W. (2003). SNIF-ACT: A model of information foraging on the World Wide Web. In P. Brusilovsky & A. Corbett & F. de Rosis (Eds.), *User Modeling 2003, 9th International Conference, UM 2003* (Vol. 2702, pp. 45-54). Johnstown, PA: Springer-Verlag.
- Reeder, R.W., P. Pirolli, and S.K. Card. Web-Eye Mapper and WebLogger: Tools for analyzing eye tracking data collected in Web-use studies. In *Human Factors in Computing Systems, CHI 01*. 2001. Seattle, WA.
- Reitman, W. R. (1964). Heuristic decision procedures, open constraints, and the structure of ill-defined problems. In M. W. Shelly & G. L. Bryan (Eds.), *Human judgements and optimality*. New York: Wiley & Sons.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.
- Simon, H. A. (1956). Rational Choice and the Structure of the Environment. *Psychological Review*, 63, 129-138.
- Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence*, 4, 181-204.
- Spool, J. M., Perfetti, C., & Brittan, D. (2004). *Designing for the scent of information*. Middleton, MA: User Interface Engineering.
- Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., & DeAngelo, T. (1999). *Web site usability*. San Francisco, CA: Morgan Kaufman.

- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Turney, P. D. (2001). Mining the Web for synonyms: PMI-IR versus LSA on TOEFL. In *Twelfth European Conference on Machine Learning, ECML 2001*, Freiburg, Germany.
- User Interface Engineering. (1999). Designing information-rich web sites. Cambridge, MA: Author.

9. FIGURE CAPTIONS

Figure 1. A schematic example of the information scent assessment subtask facing a Web user. The arrows represent associations between the words.

Figure 2. The structure of SNIF-ACT 1.0 and the User-Tracer.

Figure 3. Web Behavior Graphs for one study subject working on the ANTZ task (left) and CITY task (right) in experiment 1.

Figure 4. The links chosen by subjects and ranked by SNIF-ACT 1.0 and the Position model. The lower the rank, the more likely that the model will choose the links.

Figure 5. The mean scent scores before subjects left a Web site. The dashed line represents the overall mean scent scores of all Web pages visited by the subjects.

Figure 6. (a) A hypothetical Web page in which the information scent of links decreases linearly from 10 to 2 as the model evaluated links 1 to 5. The information scent of the links from 6 onwards stays at 2. The number in parenthesis represents the value of information scent. (b) The probability of choosing each of the competing productions when the model processes each of the link in (a) sequentially. The mean information scent of the previous pages was 10. The noise parameter t was set to 1.0. The initial utilities of all productions were set to 0. k and *GoBackCost* were both set to 5.

Figure 7. The scatter plots for the number of times links were selected in the Parcweb and Yahoo sites by subjects and by the SNIF-ACT 2.0, SNIF-ACT 1.0, and Position model.

Figure 8. The scatter plots of the number of times subjects and the model went back to the previous pages.

Figure 9. Log-Log plots of frequency against number of clicks on web pages in Yahoo and ParcWeb. In the equations, x represents $\text{Log}(\text{clicks})$ and y represents $\text{Log}(\text{frequency})$.

Figure 10. The Cumulative Distribution Frequency for the number of users on Yahoo and ParcWeb plotted against the number of clicks and the predictions by the law of surfing and SNIF-ACT 2.0.

Figure 1. A schematic example of the information scent assessment subtask facing a Web user. The arrows represent associations between the words.

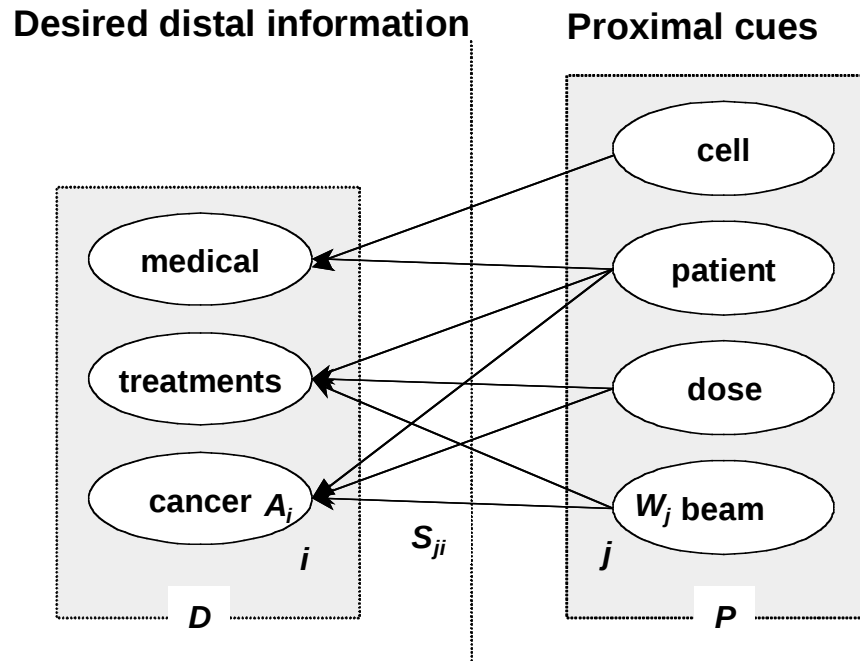


Figure 2. The structure of SNIF-ACT 1.0 and the User-Tracer.

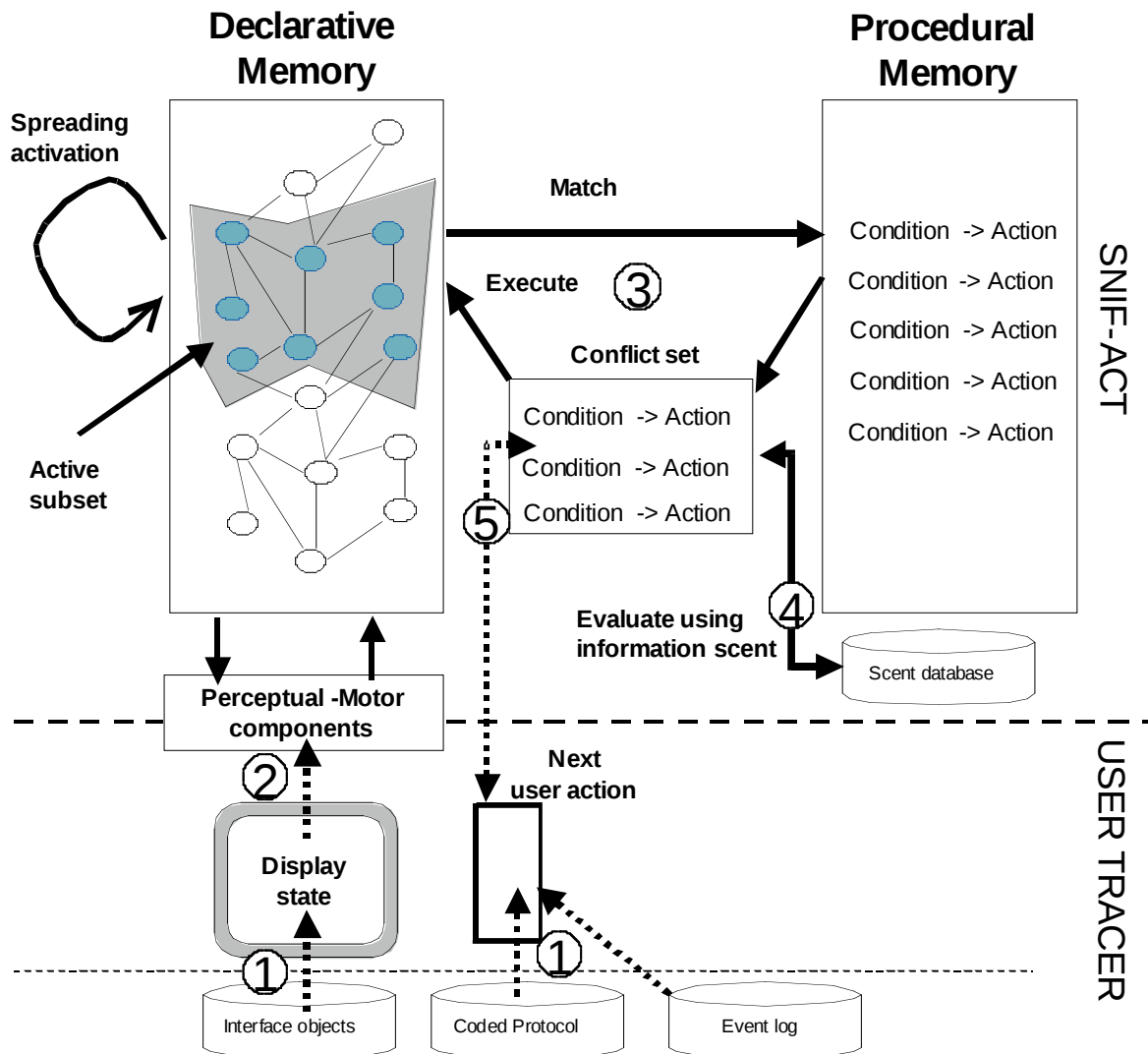


Figure 3. Web Behavior Graphs for one study subject working on the ANTZ task (left) and CITY task (right) in experiment 1.

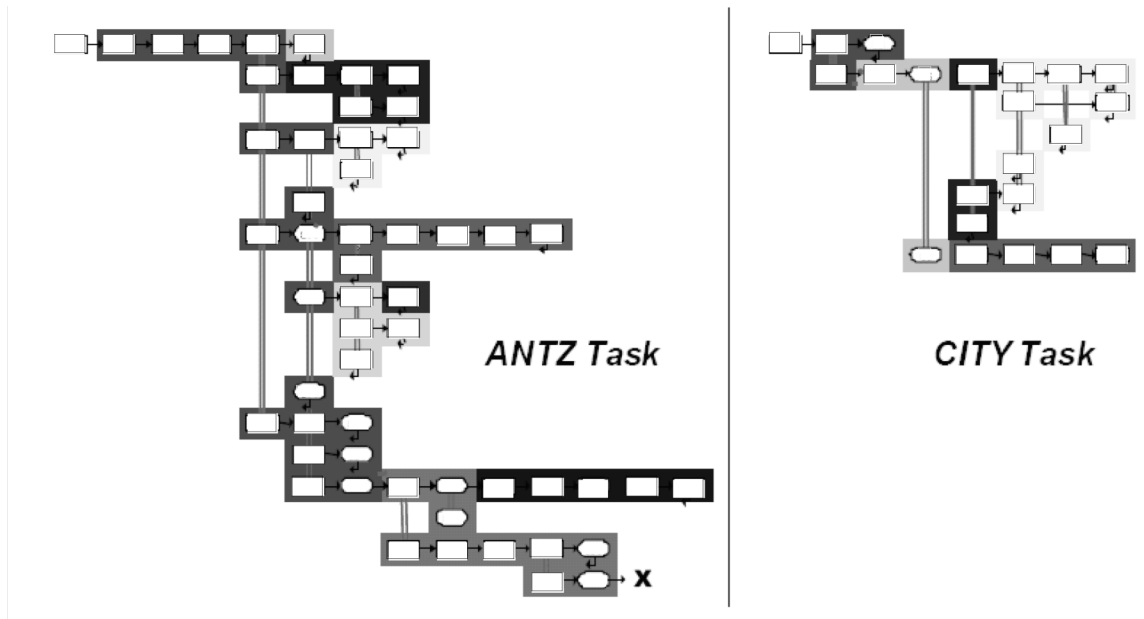


Figure 4. The links chosen by subjects and ranked by SNIF-ACT 1.0 and the Position model. The lower the rank, the more likely that the model will choose the links.

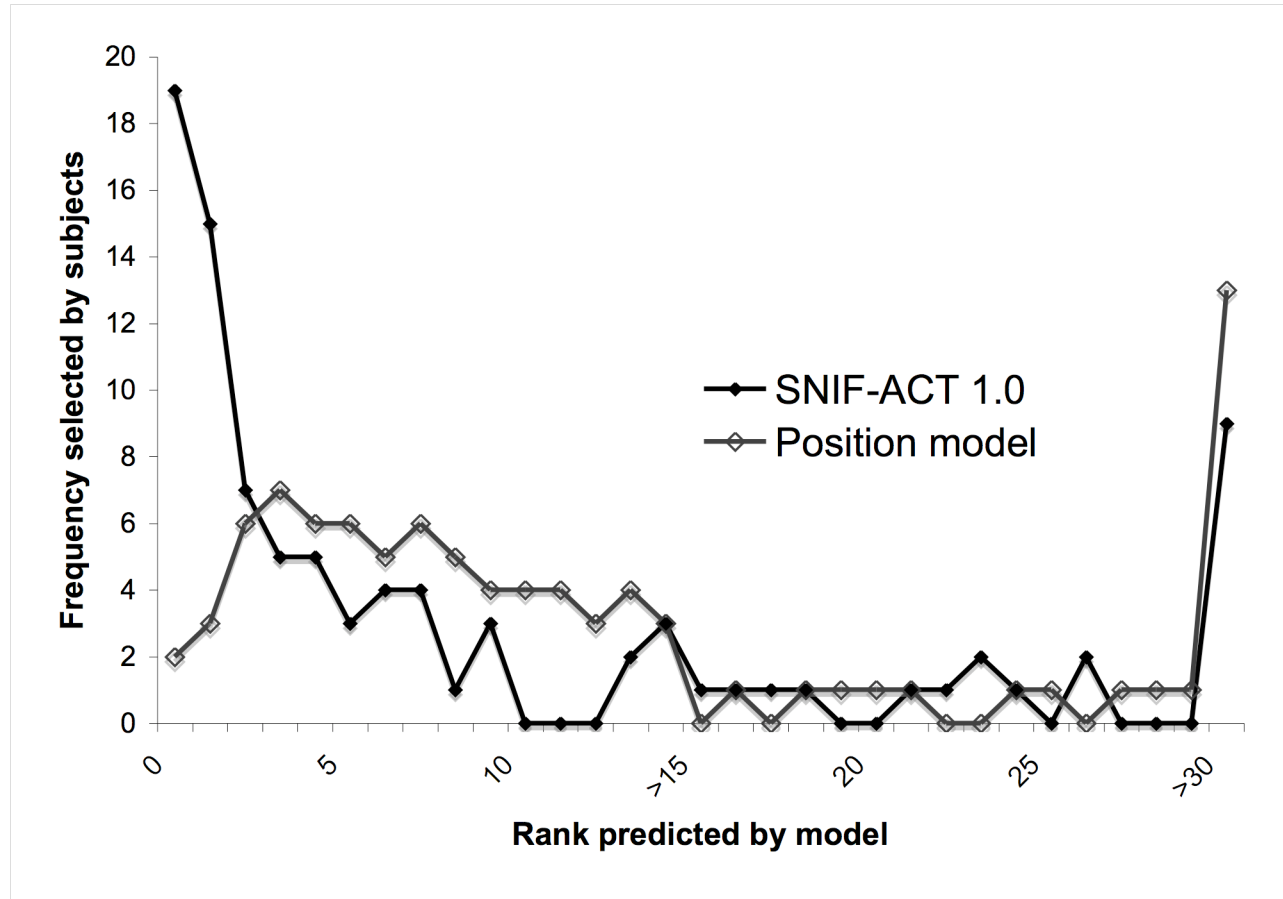


Figure 5. The mean scent scores before subjects left a Web site. The dashed line represents the overall mean scent scores of all Web pages visited by the subjects.

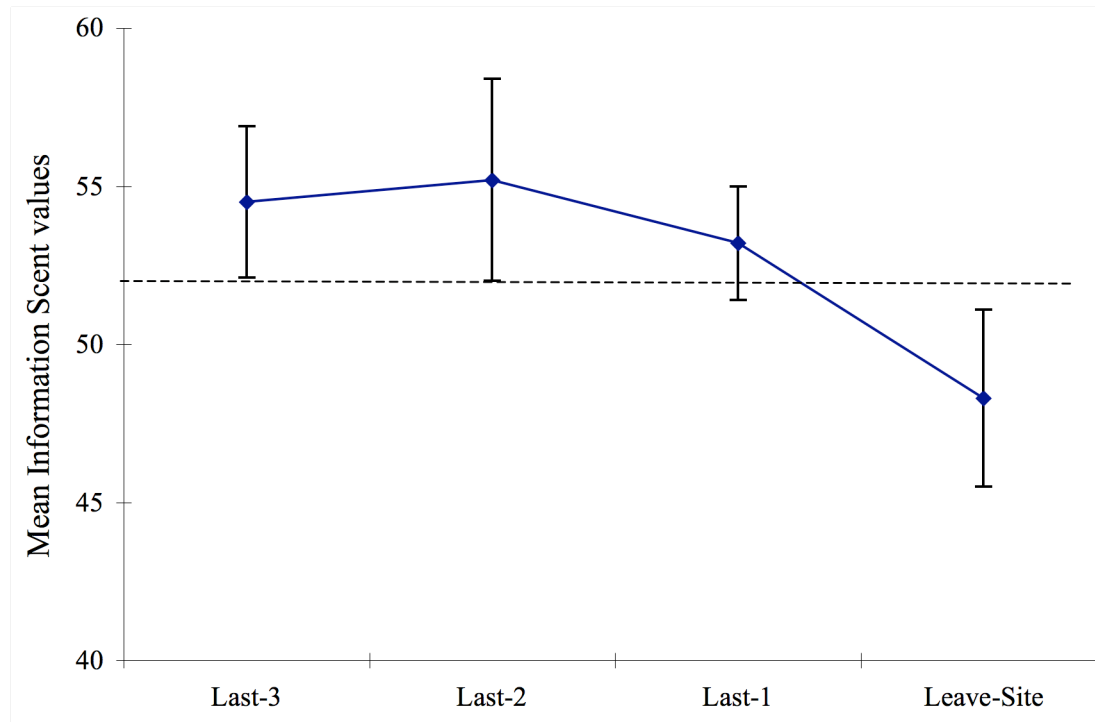


Figure 6. (a) A hypothetical Web page in which the information scent of links decreases linearly from 10 to 2 as the model evaluated links 1 to 5. The information scent of the links from 6 onwards stays at 2. The number in parenthesis represents the value of information scent. (b) The probability of choosing each of the competing productions when the model processes each of the link in (a) sequentially. The mean information scent of the previous pages was 10. The noise parameter t was set to 1.0. The initial utilities of all productions were set to 0. k and *GoBackCost* were both set to 5.

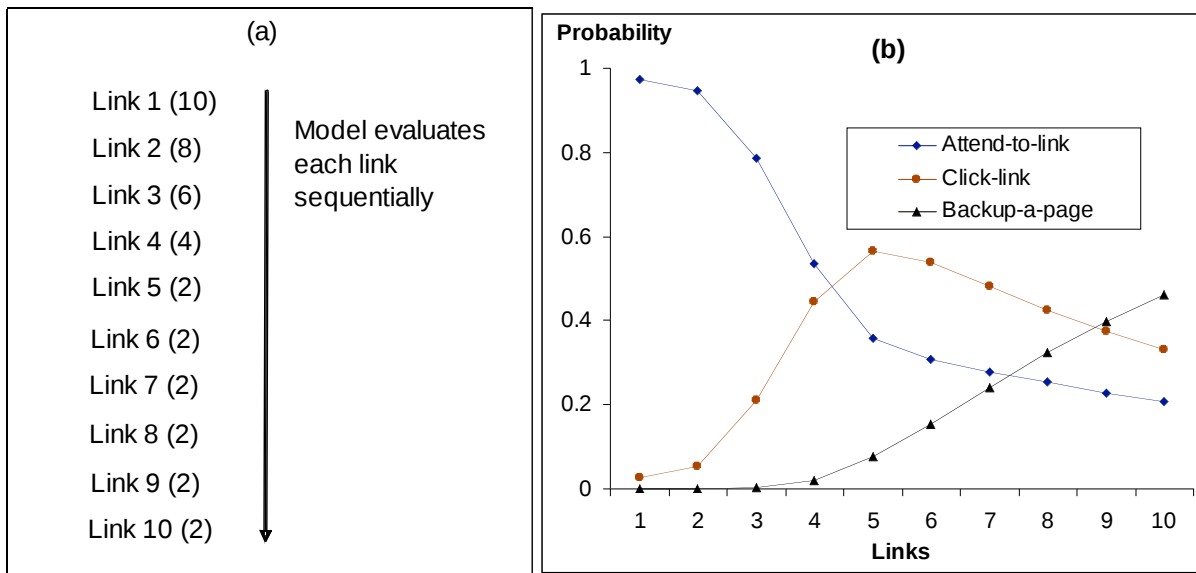


Figure 7. The scatter plots for the number of times links were selected in the Parcweb and Yahoo sites by subjects and by the SNIF-ACT 2.0, SNIF-ACT 1.0, and Position model.

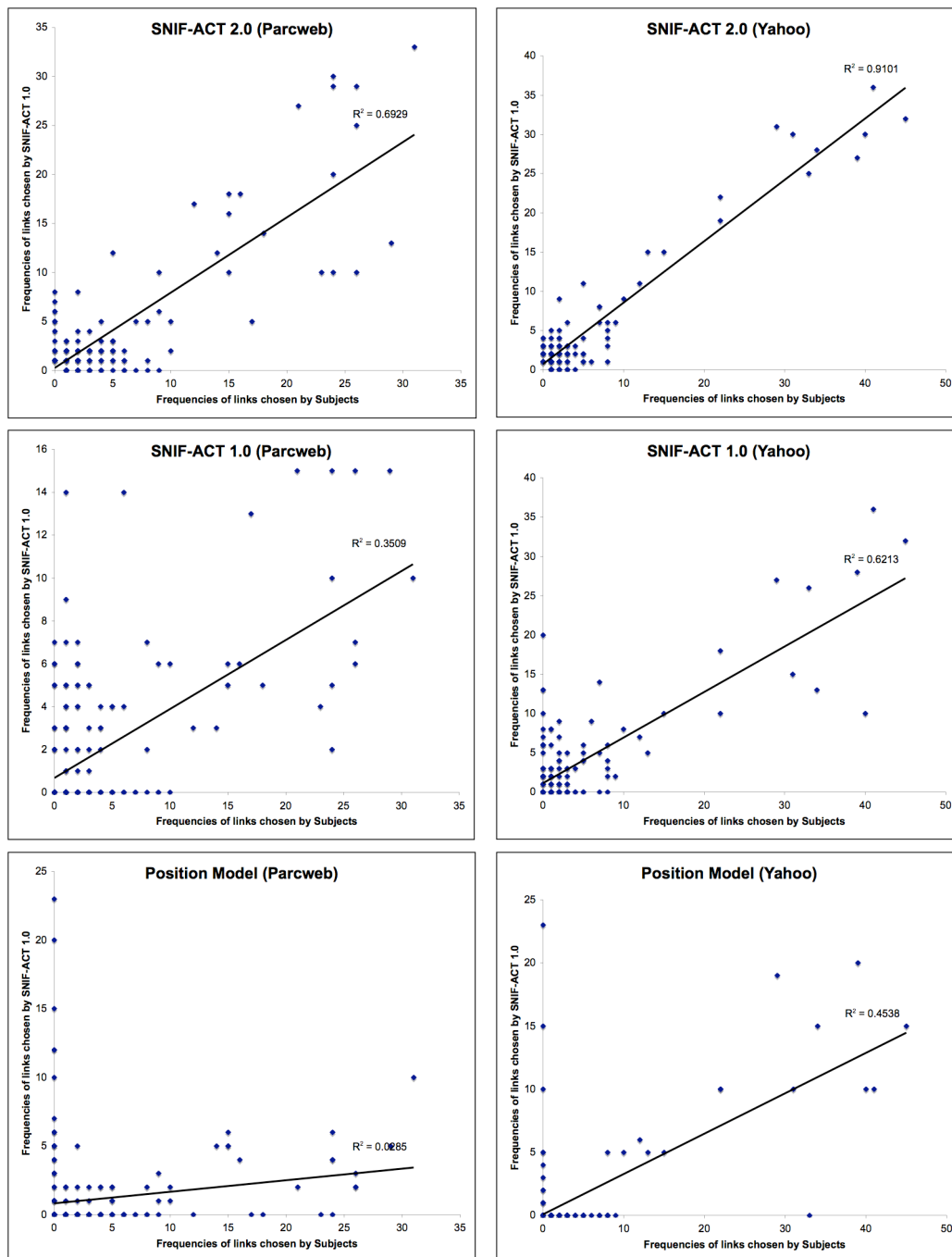


Figure 8. The scatter plots of the number of times subjects and the model went back to the previous pages.

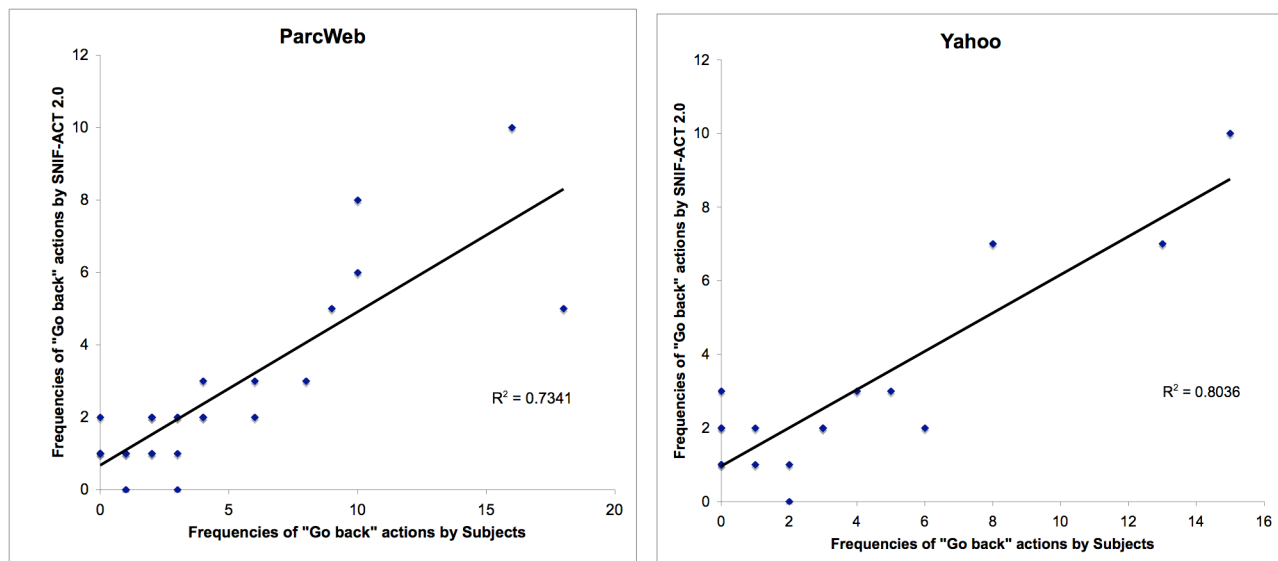
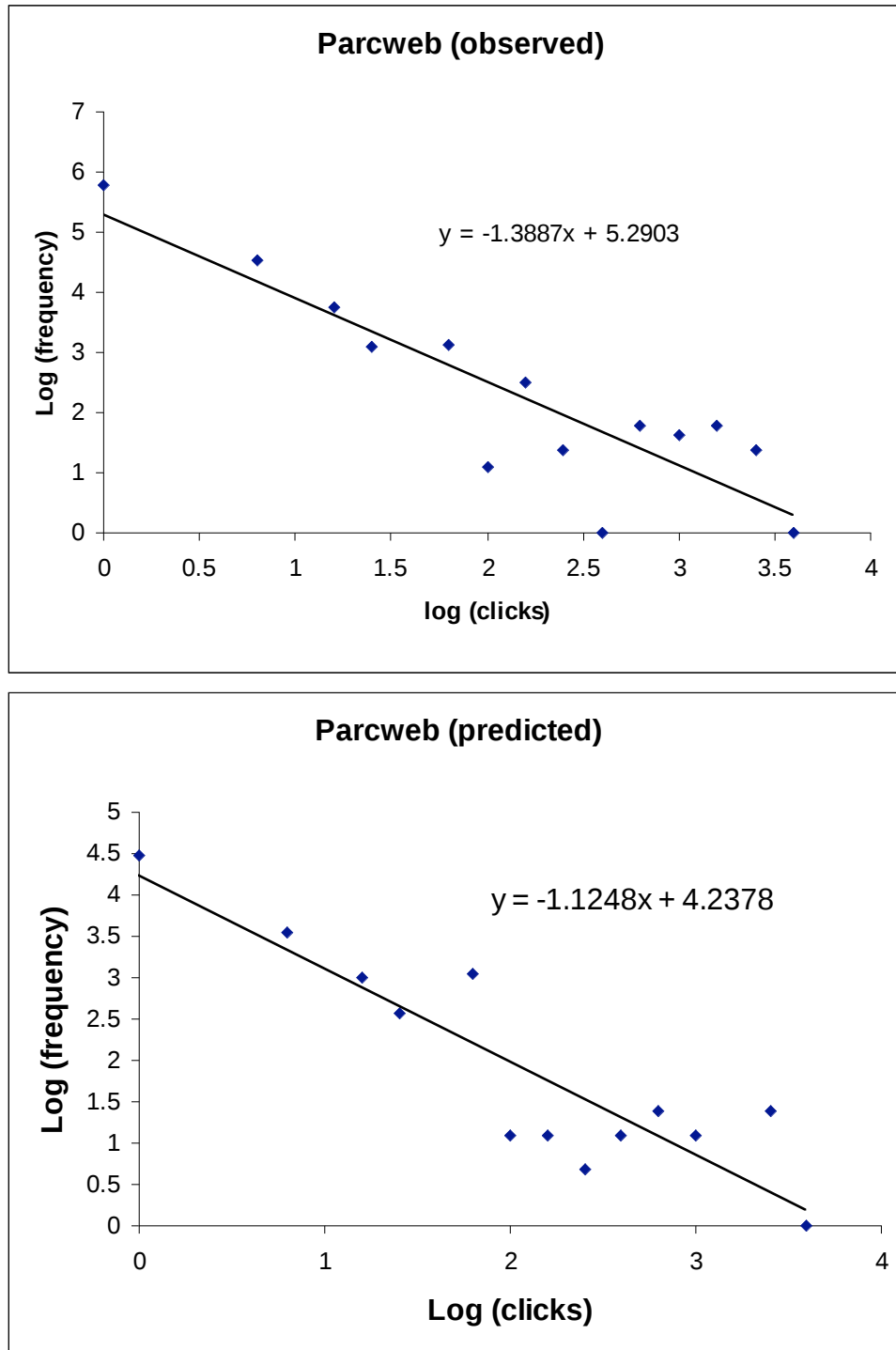


Figure 9. Log-Log plots of frequency against number of clicks on web pages in Yahoo and ParcWeb.

In the equations, x represents Log(clicks) and y represents Log(frequency).



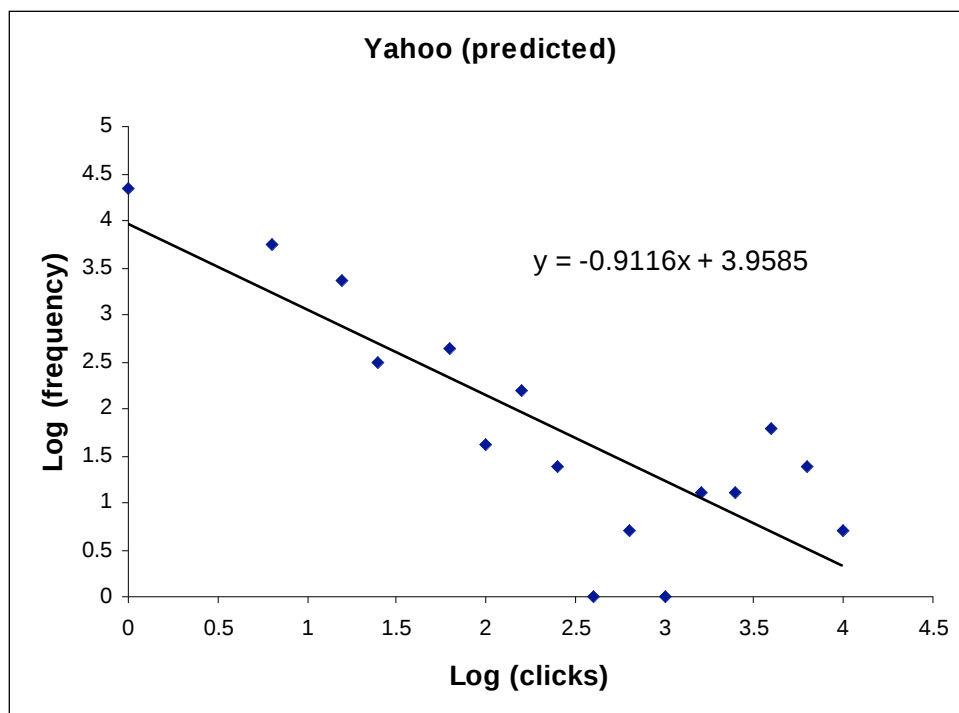
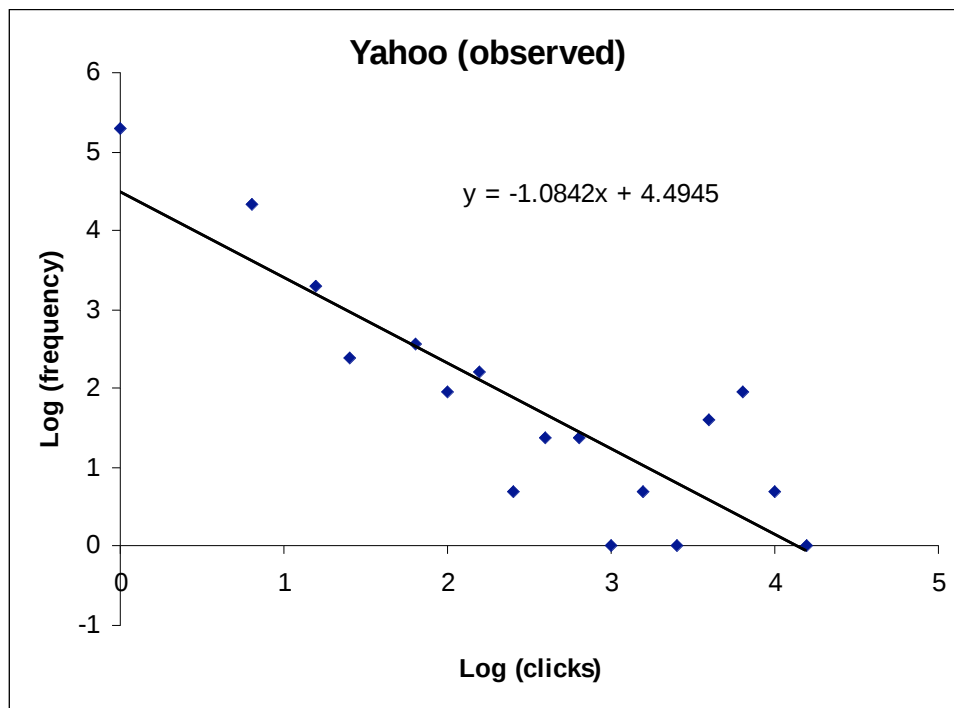
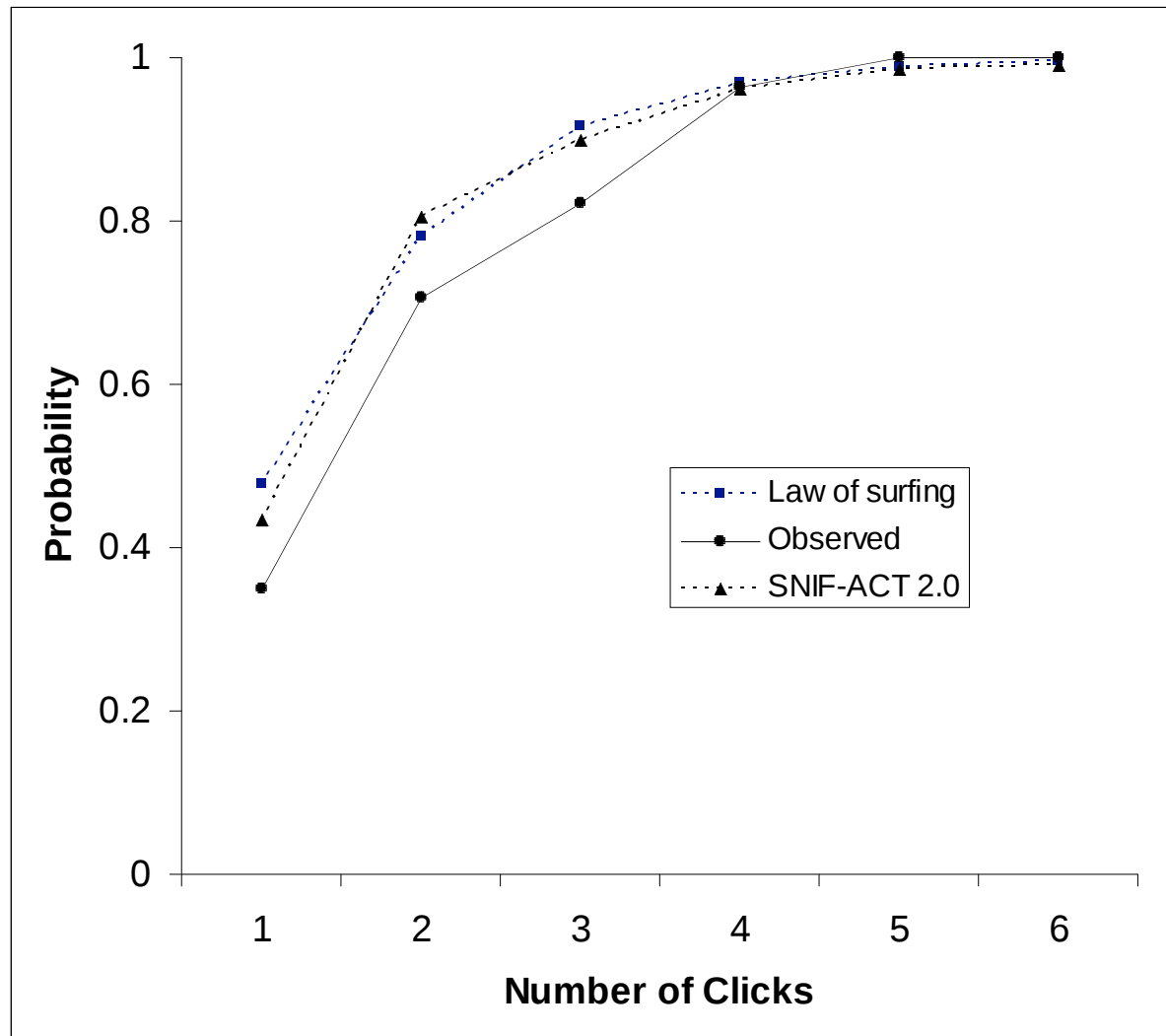


Figure 10. The Cumulative Distribution Frequency for the number of users on Yahoo and ParcWeb plotted against the number of clicks and the predictions by the law of surfing and SNIF-ACT 2.0.



10. TABLES

Table 1. Productions in SNIF-ACT 1.0 in their English equivalent forms.

Start-Process-Page: IF the goal is Goal*Start-Next-Patch & there is a task description & there is a browser & the browser is on an unprocessed page THEN Set & push a subgoal Goal*Process-Page to the goal stack
Process-Links-on-Page: IF the goal is Goal* Process-Page & there is a task description & there is a browser & there is an unprocessed link THEN Set and push a subgoal Goal*Process-Link to the goal stack
Attend-to-Link: IF the goal is Goal* Process-Link & there is a task description & there is a browser & there is an unattended link THEN Choose an unattended link and attend to it
Read-and-Evaluate-Link: IF the goal is Goal* Process-Link & there is a task description & there is a browser & the current attention is on a link THEN Read and Evaluate the link
Click-Link: IF the goal is Goal* Process-Link & there is a task description & there is a browser & there is an evaluated link & the link has the highest activation THEN Click on the link
Leave-Site: IF the goal is Goal* Process-Link & there is a task description & there is a browser & there is an evaluated link & the mean activation on page is low THEN Leave the site & pop the goal from the goal stack
Backup-a-Page: IF the goal is Goal* Process-Link & there is a task description & there is a browser & there is an evaluated link & the mean activation on page is low THEN Go back to the previous page

Table 2. An example trace of the SNIF-ACT model.

<u>Productions</u>	<u>Descriptions</u>
Use-Search-Engine fired	Model started, decided to use a search engine.
Go-To-Search-Engine fired	Retrieved address of search engine from memory.
Go-To-Site-By-Typing fired	Typed address of search engine on browser.
Start-Process-Page fired	Moved attention to new Web page.
Search-Site-using-Search-Box fired	Typed search terms in search box
Process-Links-on-Page fired	Prepared to move attention to a link on page.
Attend-to-Link fired	Moved attention to the link.
Read-and-Evaluate-Link fired	Read and evaluated the link.
Attend-to-Link fired	Moved attention to next link.
Read-and-Evaluate-Link fired	Read and evaluated the link.
Attend-to-Link fired	Moved attention to next link.
Read-and-Evaluate-Link fired	Read and evaluated the link.
Click-Link fired	Clicked on the link.
....	
Click Link	Clicked on the link.
Finish fired	Target found.

Table 3. The tasks given to subjects in Experiment 2.

<u>ParcWeb (19227 documents)</u>	
<u>Tasks</u>	
<u>1a</u>	Find the PowerPoint slides for Jan Borchers's June 3, 2002 Asteroid presentation.
<u>1b</u>	Suppose this is your first time using AmberWeb. Find some documentation that will help you figure out how to use it.
<u>2a</u>	Find out where you can download the latest DataGlyph Toolkit.
<u>2b</u>	Find some general information about the DataGlyphs project.
<u>3a</u>	What do the numerical TAP ratings mean?
<u>3b</u>	What patent databases are available for use through PARC?
<u>4a</u>	Find the 2002 Holiday Schedule
<u>4b</u>	Where can you download an expense report?
 <u>Yahoo (7484 documents)</u>	
<u>Tasks</u>	
<u>1a</u>	What is the Yahoo! Directory?
<u>1b</u>	You want Yahoo! to add your site to the Yahoo! Directory. Find some guidelines for writing a description of your site.
<u>2a</u>	You have a Yahoo! Email account. How do you save a message to your Sent Mail folder after you send it?
<u>2b</u>	You are receiving spam on your Yahoo! Email account. What can you do to make it stop?
<u>3a</u>	When is the playing season for Fantasy Football?
<u>3b</u>	In Fantasy Baseball, what is rotisserie scoring?
<u>4a</u>	You are trying to find your friend's house, and you are pretty sure you typed the right address into Yahoo! Maps, but the little red star still showed up in the wrong place. How could this have happened?
<u>4b</u>	You want to get driving directions to the airport, but you don't know the street address. How else can you get accurate directions there?

Table 4. The number of usable user sessions, Web pages visited, successes, and the number of times subjects decided to go back to previous Web page in each of the two sites.

					<u>Tasks</u>				<u>Total</u>
<u>ParcWeb</u>	<u>1a</u>	<u>2a</u>	<u>3a</u>	<u>4a</u>	<u>1b</u>	<u>2b</u>	<u>3b</u>	<u>4b</u>	
Sessions	31	27	30	33	28	29	24	30	232
Pages	124	72	120	86	350	106	107	232	1197
Successes	27	0	0	31	5	0	0	23	86
Going back	6	10	9	9	8	10	22	4	78

					<u>Tasks</u>				<u>Total</u>
<u>Yahoo</u>	<u>1a</u>	<u>2a</u>	<u>3a</u>	<u>4a</u>	<u>1b</u>	<u>2b</u>	<u>3b</u>	<u>4b</u>	
Sessions	44	47	44	44	44	43	47	45	358
Pages	104	149	164	144	216	197	260	257	1491
Successes	40	39	36	43	13	18	45	31	265
Going back	10	8	9	8	5	6	8	9	63

Table 5. The percentages of successes in each of the tasks for the subjects and the models.

					<u>Tasks</u>			
<u>ParcWeb</u>	<u>1a</u>	<u>2a</u>	<u>3a</u>	<u>4a</u>	<u>1b</u>	<u>2b</u>	<u>3b</u>	<u>4b</u>
Subject	87%	0%	0%	94%	18%	0%	0%	77%
Position	10%	0%	0%	12%	0%	0%	0%	0%
Snif-Act 1.0	61%	21%	16%	62%	8%	7%	24%	45%
Snif-Act 2.0	71%	0%	0%	63%	21%	0%	0%	51%

					<u>Tasks</u>			
<u>Yahoo</u>	<u>1a</u>	<u>2a</u>	<u>3a</u>	<u>4a</u>	<u>1b</u>	<u>2b</u>	<u>3b</u>	<u>4b</u>
Subject	91%	83%	82%	98%	30%	42%	96%	69%
Position	13%	9%	2%	21%	2%	6%	15%	7%
Snif-Act 1.0	53%	76%	78%	82%	21%	37%	46%	53%
Snif-Act 2.0	89%	79%	76%	88%	16%	24%	78%	45%

APPENDIX A: THE RANDOM UTILITY MODEL OF LINK CHOICE

Consider a person facing a Web page with a choice set of links L consisting of j alternatives. Suppose the person chooses alternative k from L . If rational behavior is assumed, revealed preference implies that $U_k \geq U_j$ for all j in L . The probability of this event occurring can be represented as $P_k = \text{Prob}(U_k \geq U_j \text{ for all } j \text{ in } L)$.

In the random utility model, utilities are assumed to consist of two parts, one is deterministic and one is stochastic, thus the utility of link k can be represented as

$$U_k = V_k + \varepsilon_k$$

$$\begin{aligned} \text{Thus, } P_k &= \text{Prob}(U_k \geq U_j \text{ for all } j \text{ in } L) \\ &= \text{Prob}(V_k - V_j \geq \varepsilon_j - \varepsilon_k \text{ for all } j \text{ in } L) \end{aligned}$$

To determine P_k (the probability that a person will choose link k), one needs to specify the distribution for ε . McFadden (1974) shows that if we allow the assumption that ε follows one of the popular extreme value distribution called a double exponential distribution, i.e., $\text{Prob}(\varepsilon_k < t) = \exp[-\exp(-t/b)]$, then one can obtain the conflict resolution equation as:

$$P_k = \frac{\exp(V_k / \tau)}{\sum_{j \in L} \exp(V_j / \tau)}, \text{ where } \tau = \sqrt{2} b.$$

The assumption of the double exponential distribution for the error term thus allows an elegant closed form equation for the probability of selecting a link from a set. The distribution corresponds to the limiting distribution of the maximum value in a set of N elements as N approaches infinity.

APPENDIX B: A RATIONAL ANALYSIS OF LINK EVALUATION AND SELECTION

The analysis below is based on the rational analysis in chapter 5 of Anderson (1990). The analysis aims at providing a rational basis for the utility calculations of the productions in the SNIF-ACT 2.0 model. The goal of the rational analysis is to derive the adaptive mechanism for the action evaluation and selection process as links are sequentially processed. The analysis is based on a Bayesian framework in which the user is gathering data from the sequential evaluation of links on a web page. We define:

X = variable that measures the closeness to the target

S = binary variable that describes whether the link will lead to the target page

R = probability that the target information can be found

r = the event that the target information exists

Given the definitions, it immediately follows that

$$\Pr(S = 1|r) = R, \quad (\text{A.1})$$

and

$$\Pr(S = 0|r) = 1 - R; \quad (\text{A.2})$$

we also have, by Bayes Theorem:

$$\Pr(S, X|r) = \Pr(X|S, r) \Pr(S|r). \quad (\text{A.3})$$

Since the major assumption of the Information Foraging Theory is that information scent (IS) directly measures the closeness to the target, we define:

$$P(X | S = 1, r) = K \sum_{j=0}^n \frac{\Gamma(\alpha + j)}{\Gamma(\alpha + n)} IS(j) \quad (\text{A.4})$$

where $IS(j)$ represents the information scent of link j , and K and α are constant parameters of the Equation A.4. The above Equation A.4 assumes that the measure of closeness is a hyperbolically discounted sum of the information scents of the links encountered in the past. The use of a hyperbolic discount function has been validated in a number of studies

in human preferences (e.g., Loewenstein & Prelec, 1991; Ainslie & Haslam, 1992; Mazur, 2001).

We treat this problem as one of sampling the random variables (X, S) from a Bernoulli distribution, with R equivalent to the parameter to be estimated for the distribution. The appropriate Bayesian conjugate distribution convenient for use in updating estimates of R from samples of a Bernoulli random variable is the beta distribution. That is, we assume a prior beta distribution for R , and the user will use the observed information scent of the links on a web page to update a posterior beta distribution of R . We take R to follow a beta distribution with parameters a and b . After the user has experienced a sequence of links on a Web page, represented as

$$L_n = ((X_1, S_1), (X_2, S_2) \dots (X_n, S_n))$$

where each pair (X_i, S_i) describes the closeness to the target and whether the link leads to the target page. Since the prior of R is a beta distribution, the posterior distribution $\Pr(R|L_n)$ is also a beta, and the new parameters can be shown to be

$$a_{new} = a + \sum S_i \quad (A.5)$$

and

$$b_{new} = b + \sum (1 - S_i) \quad (A.6)$$

as its parameters. The posterior predictive distribution for S and X given L_n can be computed as:

$$\Pr(S_{n+1}, X_{n+1} | L_n) = \int \Pr(S_{n+1}, X_{n+1} | R) \Pr(R | L_n) dR \quad (A.7)$$

In our case, our interest mainly lies on the posterior predictive probability that the user can find the target, i.e., $\Pr(S_{n+1} = 1, X_{n+1} | L_n)$, which can be computed as:

$$\Pr(S_{n+1} = 1, X_{n+1} | L_n) = \left(K \sum_{j=0}^n \frac{\Gamma(\alpha + j)}{\Gamma(\alpha + n)} IS(j) \right) \left(\frac{a + \sum S_i}{a + b + n} \right) \quad (A.8)$$

If the user is considering links sequentially on a Web page before the target is found, we have $\sum S_i = 0$. To reduce the number of parameters, we set $\alpha = a$, $K = 1/a$. and assume that

$b = 0$. We now only have one parameter a , which represents the prior number of successes in finding the target information on the web. The equation can then be reduced to:

$$\Pr(S_{n+1} = 1, X_{n+1} | L_n) = \sum_{j=0}^n \frac{\Gamma(a+j)}{\Gamma(a+n+1)} IS(j) = U(n) \quad (\text{A.9})$$

In the model, the above probability is calculated to approximate the utilities of the productions read-next-link and click-link. Putting the above equation in a recursive form, we have:

$$U(n) = \frac{U(n-1) + IS(n)}{a+n} \quad (\text{A.10})$$

In the equation specified in the text, we set $a = 1$ for the read-next-link production; and $a = 1 + k$ for the click-link production. By setting the value of a for click-link to a higher value, we assume that in general, following a link is more likely to lead to the target page than attending to the next link on the same web page. k is a free parameter that we used to fit the data.

APPENDIX C: THE LAW OF SURFING

The law of surfing (Huberman, Pirolli, Pitkow, & Lukose, 1998) was derived to describe emergent aggregate web navigation behavior. We are interested to see (1) if the data sets we collected also exhibit the same properties as predicted by the law of surfing (LoS), and (2) whether SNIF-ACT 2.0, a model aims at explaining fine-grained dynamic user-Web interactions, will exhibit the same emergent properties at the aggregate level. The LoS is based on the notion that Web surfing can be modeled as a Weiner process with a random (positive) drift parameter μ and with noise σ^2 . Specifically, the utility of a page X_t to be visited at time t to the utility of a currently viewed page X_{t-1} at time $t - 1$ is calculated as

$$U(X_t) = U(X_{t-1}) + \varepsilon_t \quad (1)$$

where ε_t is a random variable from a Gaussian distribution with mean μ and variance σ^2 . It is assumed that this process starts in some initial state X_0 , and terminates when some threshold utility U is encountered. The distribution of first passage times (i.e, in our case, the number of clicks on a web site before the user leaves, or the “depth”) for this process is characterized by an Inverse Gaussian Distribution (IGD) which is usually presented as

$$\begin{aligned} f(t; v; \lambda) &= \sqrt{\frac{\lambda}{2\pi t^3}} e^{-\frac{\lambda(t-v)^2}{2v^2t}}, t > 0 \\ &= \sqrt{\frac{\lambda}{2\pi}} t^{-3/2} e^{-\frac{\lambda(t-v)^2}{2v^2t}}, t > 0 \end{aligned} \quad (2)$$

where $E[t] = v$ and $Var[t] = \frac{v^3}{\lambda}$. An interesting implication of the LoS can be obtained by taking logarithms on both sides of (2), which yields

$$\log f(t) = -\frac{3}{2}\log t - \frac{\lambda(t-v)^2}{2v^2t} - \log\left(\sqrt{\frac{\lambda}{2\pi}}\right) \quad (3)$$

The equation suggests that a log-log plot will show a straight line whose slope approximates -3/2 for small values of t . Figure 9 shows the log-log plot of the observed and predicted frequency and the number of clicks in the Yahoo and ParcWeb web sites.

We can see that in general both the observed and predicted data by SNIF-ACT 2.0 are consistent with the properties predicted by the LoS.

----- INSERT Figure 9 ABOUT HERE -----

The LoS also allows precise predictions on the probability that a user will leave a web site as the user is navigating on the web site. Figure 10 shows the cumulative distribution frequency (CDF) of the predictions by LoS and SNIF-ACT 2.0. The figure also shows the data collected from Yahoo and ParcWeb, with mean of 2.31 clicks and variance of 1.35 before users stopped clicking forward (i.e., either go back, type in a different URL, etc.). The match between the predictions between the LoS and SNIF-ACT 2.0 are extremely good ($R^2=0.993$), and the match between the observed and LoS ($R^2=0.984$) and that between the observed and SNIF-ACT 2.0 ($R^2=0.976$) are also good. The good match between SNIF-ACT 2.0 and LoS in Figure 9 and Figure 10 is striking. SNIF-ACT and LoS were derived based on very different assumption of human behavior and contents of Web sites. LoS was derived based on minimal assumption of human behavior (the IGD) and was insensitive to specific contents of Web pages. The value of LoS is its predictive power in long-term aggregate behavior in very large information structures. On the other hand, SNIF-ACT was derived from a rational analysis of link selection and assumption of how a single user may dynamically make decisions based on specific contents of Web sites. The good match between the two model suggests that the long-term expected behavior of SNIF-ACT is consistent with the predictions by LoS.

----- INSERT Figure 10 ABOUT HERE -----