Cognitive Architecture for Website Design and Usability Evaluation: Comprehension and Information Scent in Performing by Exploration

Muneo Kitajima

National Institute of Advanced Industrial Science and Technology (AIST)
1-1-1, Higashi, Tsukuba, Ibaraki 305-8566 Japan
kitajima@ni.aist.go.jp

Marilyn Hughes Blackmon and Peter G. Polson

Institute of Cognitive Science,
University of Colorado
Boulder, CO 80309-0344 USA
{ppolson, blackmon}@psych.colorado.edu

Abstract

CoLiDeS is a model of how people navigate a complex website to find information and is the youngest member in a family of models for human-computer interaction situations where users rely primarily on skilled reading and action planning as the core cognitive processes. All models in this family are based on Kintsch's (1998) construction-integration (C-I) cognitive architecture. This paper describes CoLiDeS in relation to C-I, highlighting features of CoLiDeS that distinguish it from SNIF-ACT, a competing model of website navigation based on the ACT-R cognitive architecture combined with Information Foraging theory. Highlighting differences between CoLiDeS and SNIF-ACT yields insights into the consequences of selecting one cognitive architecture over another, helping us uncover the future potential for cognitive architectures in HCI. Controlled laboratory studies of the Cognitive Walkthrough for the Web (CWW), an engineering approximation of CoLiDeS, offer strong empirical evidence for the psychological validity and reliability of CoLiDeS and suggest a promising future for C-I in HCI.

1 Introduction

A major advantage of cognitive architectures is their comprehensiveness and versatility to cover a wide range of human behaviors. We have developed a family of models derived from the Construction-Integration (C-I) cognitive architecture (Kintsch, 1998) to account for various aspects of human-computer interaction (HCI), including skilled use of applications hosted on systems with graphical user interfaces (Kitajima and Polson, 1995), performing novel tasks by exploration (Kitajima and Polson, 1997), and browsing a website to find information relevant to a user's goal (Kitajima, Blackmon, and Polson, 2000). The C-I architecture was originally developed to explain skilled reading comprehension and later extended to action planning by Mannes and Kintsch (1991).

Our papers about this family of models show how useful the C-I cognitive architecture is for modeling such important HCI processes as comprehending task instructions, forming goals, and navigating a complex website. Navigating a website requires comprehension of texts or text fragments (e.g., headings, hyperlinks, and content), and scanning and reading webpages requires comprehension of visuospatial information, not just information communicated by texts. Users must be able to understand webpage layouts analogous to the way that skilled readers understand page layout and headline conventions of newspapers, the organizational structure that is communicated by headings, subheadings, and sidebars of a chapter or article, and by the table of contents in a book.

In this paper we will briefly describe C-I cognitive architecture and its extension to action planning, and the family of HCI models based on C-I. We then offer a detailed description of the CoLiDeS model of web navigation that is based on C-I architecture, calling special attention to features of CoLiDeS that distinguish it from SNIF-ACT (Pirolli et al., 2002; Pirolli & Fu, 2003; Pirolli, 2004), a competing model of website navigation derived from the ACT-R cognitive architecture (Anderson & Lebiere, 1998; Anderson et al., 2004) combined with the Information Foraging (IF) theory (Pirolli & Card, 1997). After briefly reviewing the past and present uses of C-I cognitive architecture in HCI, we look at the future potential of C-I in HCI and whether HCI research can contribute to the further development of C-I. We explain the reliance of CoLiDeS on Latent Semantic Analysis (LSA) for computing semantic similarity but accentuate the distinctive way CoLiDeS uses LSA to assess whether users have adequate background knowledge for comprehending the links and headings on a webpage. Finally, we discuss the empirical verification of CoLiDeS, drawing on laboratory experiments we conducted to test the Cognitive Walkthrough for the Web (CWW) (Blackmon et al., 2002, 2003, 2005), an engineering approximation of CoLiDeS..

2 Simulated Users: Description of CoLiDeS as Built on C-I Cognitive Architecture

2.1 Human-Computer Interaction Viewed as Comprehension Processes

In C-I cognitive architecture, text comprehension is a constraint-satisfaction process that utilizes a massive amount of background knowledge stored in long-term memory in order to identify the meaning of a sentence that is consistent with the current context. In C-I cognitive architecture, action planning uses the same constraint-satisfaction process to select a correct action consistent with the current context and with the user's current goal.

In HCI, performing an action entails an object to be acted on. And thus, all members of the family of related models (Kitajima and Polson, 1995, 1997) view action planning as a two-stage process. In the first stage many objects in a display (e.g., hypertext links, pull-down menus, text entry boxes, etc.) compete for attention as possible targets for action, so the first stage is *an attention process* that ends by selecting to attend to a group of display objects that is similar to the user's goal. In the second stage, the user's attention is focused on a small number of objects in the attended-to group, but there are many possible actions on each object (e.g., type, click, double-click). The second stage is thus *an action-selection process* that selects an object-action pair whose description is perceived most similar to the user's goal and consistent with the properties of the object (e.g., press and hold on a menu label, click on a hypertext link, etc.)..

The C-I cognitive architecture assumes that constructing meaningful representations of texts, of objects on a display, or of object-action pairs are all similar processes analogous to reading comprehension. A text parser generates the textbase, i.e., a low level semantic representation of a sentence or meaningful text fragment. The textbase is elaborated with information retrieved from long-term memory. Some of the retrieved information is irrelevant to the current context. A spreading activation, constraint satisfaction process, retains relevant information and eliminates irrelevant information, thereby generating a coherent representation of the meaning of the input text..

Comprehension of displays and action planning are other processes analogous to reading comprehension. A visual parser generates a low level propositional representation of objects on a display. The attention process elaborates propositional representations with information retrieved from long-term memory, including properties of the objects. During the action-selection process possible actions for an object of a specific type – e.g., pull-down menus or text entry boxes – are also retrieved from long-term memory. The constraint satisfaction process causes a model to attend to the objects on the screen and to select the object-action pair most similar to the user's current goal.

Kitajima and Polson (1995) developed the basic foundations for C-I models of users interacting with applications hosted on systems with a graphical user interface (GUI) and on the World Wide Web. The paper described the propositional representations of displays and the object-action pairs, and the two-stage action planning process outlined in a preceding paragraph. They also developed techniques for simulating interactions with mice, windows, menus, dialog boxes, and other kinds of widgets found in screen displays of applications. The C-I model presented in the paper simulated a skilled user of a Macintosh application, Cricket Graph, for drawing data graphs.

LICAI (Kitajima and Polson, 1997), an acronym for LInked model of Comprehension-based Action planning and Instruction taking, simulates a new user of an application following written instructions to perform a task. Simulation of comprehending task instructions derived straight from Kintsch's (1988, 1998) C-I models of word and algebra story problems. They simulated users reading written instructions and performing novel tasks by exploration. They also showed that LICAI accounts for the results of other experiments on performing by exploration (Engelbeck, 1986; Polson & Lewis, 1990; Franzke, 1994, 1995), including *the label-following strategy* – a form of *pure forward search*, one of the problem-solving heuristics most frequently used by users at all levels of expertise. To guide search during exploration, label-following uses the overlap between users' goals and labels on menus, buttons, and other objects. Users act on interface objects with overlapping labels to try to discover correct action sequences.

CoLiDeS, an acronym for Comprehension-based Linked-model of Deliberate Search, extends LICAI in order to simulate users navigating within an informational website for a content page containing the information that the visitor wants to retrieve. CoLiDeS, like LICAI, simulates the label-following strategy, but there is an important difference between the displays LICAI must comprehend and those that CoLiDeS must comprehend. LICAI faces application displays that require knowledge about WIMP interfaces, and for LICAI it is obvious to attend to the region of the display where the next action takes place, i.e., menu bar, pulldown menu, dialog box, and so on. In contrast, it is not obvious to CoLiDeS which region of the display should become the focus of its attention.

The webpages CoLiDeS encounters require knowledge about webpage layout, a complex mixture of conventions from print media, hypermedia, and GUI applications. Extending C-I models of human-computer interaction to the web involved developing a more robust, realistic model of the attention process during web navigation, because clicking a link confronts the user with a new page containing many targets for action. CoLiDeS addressed this challenge by adding a second pair of C-I cycles. The two pairs of construction-integration cycles go through the attention and action-selection processes twice. The first pair of construction-integration cycles parses the webpage into subregions and ends by selecting an attention action, i.e., the first pair of C-I cycles ends by focusing attention on a subregion of a webpage most similar to the user's goal. The second pair of C-I cycles ends by selecting a possible action (e.g., clicking on a link) on an object from the attended-to subregion. More specifically, the second round identifies targets of action within the attended-to subregion and then an action-selection process selects and acts on a specific widget (e.g., a link) from the attended-to subregion. This second pair of construction-integration is identical to object selection and object-action pair selection in our previous models of HCI.

2.2 Current Working Version of CoLiDeS

In this section we depict a CoLiDeS simulation in relation to the concrete example illustrated in Figure 1. We first explain the goal formation mechanisms of CoLiDeS. Then we provide more detailed information about the two pairs of C-I cycles that provide the robust model of attention processes in web navigation. Then we describe the complex way that CoLiDeS computes information scent and level of background knowledge, clarifying how CoLiDeS draws on multiple sources to compute activation values. Finally, we outline the learning mechanisms of CoLiDeS. As we will show, attention to subregions and background knowledge level are crucial for differentiating CoLiDeS from the version of SNIF-ACT as described by Pirolli and Fu (2003).

2.2.1 Goal Formation

CoLiDeS creates two types of subgoals, *navigation subgoals* for how to find a webpage, and *content subgoals* for acquiring target information. Suppose, for example, that a visitor wanted to know about type 2 diabetes and started searching at WebMD, a medical website. CoLiDeS assumes that the visitor creates a goal consisting of a navigation subgoal and a content subgoal. For the CoLiDeS simulation in Figure 1, the goal contains both a content subgoal and a navigation subgoal: "I want to learn about type 2 diabetes because I am having the common symptoms of diabetes; increased thirst, frequent urination, and increased hunger (content subgoal). I will search the information by using the site navigation bar (navigation subgoal)." Either or both subgoals can be active at any given moment in time.

2.2.2 Attention Phase: First Pair of Construction-Integration Cycles

During the attention phase the user parses the webpage into subregions. Given a webpage, CoLiDeS assigns an identity at the layout level to each element on the page by applying knowledge about conventions to render a webpage. At the core is knowledge of print conventions necessary to comprehend print media, including the knowledge for identifying headings, content text, photos, their captions, advertisements, and so on. Experience with webpages adds knowledge of application and Internet conventions, including the knowledge for identifying site logo, navigation bar, text links, and traditional GUI components, such as search windows, navigation tabs, navigation control buttons, and so on. These pieces of knowledge are activated both through visuospatial bottom-up processes and knowledge-driven top-down processes. Identities of the webpage elements are determined on the basis of the degree of consistency of the recognized features with those specified in the knowledge.

Figure 1 shows an example representation that CoLiDeS would create for the "Diseases, Conditions and Health Topics" page of WebMD Health. Assuming that background knowledge is adequate to comprehend the subregion headings, CoLiDeS selects the subregion whose description is most similar to the description of the user's goal. The goal consists of both a navigation subgoal and a content subgoal, and thus, when the page layout conforms to the user's expectation, a subregion consistent with the description of the navigation subgoal would be selected. Figure 1 shows that CoLiDeS selected the site navigation bar as the end result of the attention phase.

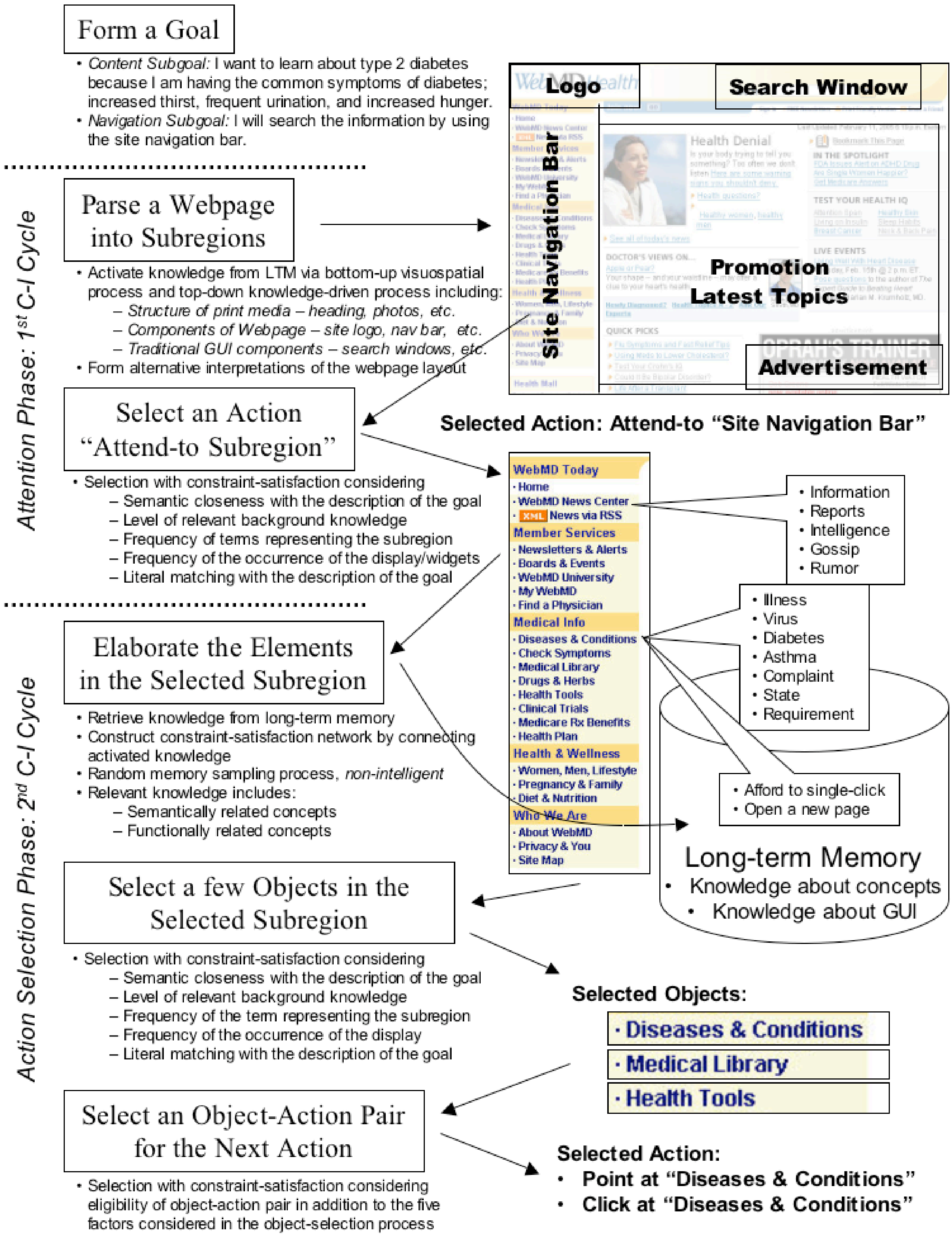

Figure 1: CoLiDeS simulation of a visitor searching WebMD (http://my.webmd.com) for information on type 2 diabetes and selecting the action to click the link "Diseases & Conditions"

2.2.3 Action Selection Phase: Second Pair of Construction-Integration Cycles

CoLiDeS elaborates the elements in the selected subregion by retrieving relevant knowledge from long-term memory. This elaboration process is a random memory sampling process cued by the representation of each object in the selected subregion. The probability of the cue retrieving an element from long-term memory is proportional to the degree of their mutual relatedness. An object is associated semantically with other concepts and functionally with interface conventions. For example, the text-link object “Diseases & Conditions” is semantically associated with such concepts as Illness, Diabetes, or Complaint, and functionally associated with Afford to single-click, Open a new page, Change color when selected, etc. The elaboration is a non-intelligent process that activates as much knowledge as possible within a given time span. Semantic elaboration is performed by referring to a semantic space that the simulated users would have and selecting terms close to the concept of the elaborated object. Functional elaboration is simulated by assuming a plausible set of knowledge typical users with GUI applications would have.

During object selection a few screen objects are selected as the candidates for next action. In Figure 1, Diseases & Conditions, Medical Library, and Health Tools, all under the heading Medical Info would be selected as the candidates for the next action. Object-action selection occurs next after object-selection. CoLiDeS retrieves representations of actions associated with the application from long-term memory and combines them with the selected objects to create executable actions. In the case of web browser, the actions associated with text links include Point and Click. Point is the prerequisite for Click but it might also cause ALT text to appear. Click is to open the linked page. The representation of the Click action is consistent with the user content goal, and thus the action Point at Diseases & Conditions is selected and then Click at Diseases & Conditions is selected.

2.2.4 How CoLiDeS Calculates Information Scent and Assesses Background Knowledge

The key notion in this section is that CoLiDeS integrates information from multiple sources to generate activation values, the equivalent of information scent in the C-I cognitive architecture. *Information scent* is a key concept for explaining users’ navigation behaviors on the Web (e.g., Blackmon et al., 2002, 2003, 2005; Chi et al., 2001, 2003; Furnas, 1997; Pirolli, 2004; Pirolli & Card, 1999; Pirolli & Fu, 2003). The metaphor evokes the image of a user searching for information by following a trail, repeatedly pursuing whatever object currently provides the highest degree of scent. Information scent is proportional to the activation level of the screen object (e.g., link), and the operational definition of scent-following in CoLiDeS is selecting the screen object with the highest activation level.

In C-I propositional networks, representations of subgoals, representations of the screen, and pieces of knowledge retrieved from long-term memory are represented as nodes with some strength level and must be interconnected with weighted links. In CoLiDeS, five independent factors combine to determine the composite information scent between the user’s goal and the screen object (i.e., heading or link text). CoLiDeS uses LSA measures as engineering approximations for the first three of these five independent factors (see below, Section 3 on LSA):

- The degree of semantic similarity between the user’s goal and the heading/link text (LSA cosine)
- Whether there is an adequate level of relevant background knowledge for successfully elaborating the heading or link (minimum LSA term vector length in the selected semantic space estimates the amount of associated knowledge in the semantic space, e.g., amount of knowledge about diseases and conditions)
- Whether a word used in the heading or link text is a low-frequency term in the user’s background knowledge (minimum LSA word frequency in the selected semantic space, a parameter that is especially crucial for very low frequency terms and zero-frequency terms typically ignored by LSA and by people)
- The frequency with which the user has encountered the screen object/widget or specific heading/link (screen elements on frequently navigated paths are more likely to be selected, e.g., a frequent user of websites with site navigation tab menus would have a propensity to navigate a website using the site tab menu, and, analogously, a person who had often used site search engines would be more apt to focus on the search window than someone who had previously located information primarily by browsing)
- Whether there is a literal matching, partial or complete, between the user’s target goal and a screen object (e.g., looking for information about Type 2 Diabetes and seeing a link labeled “Type 2 Diabetes”).

A full running simulation of CoLiDeS will integrate all the above five factors into a single activation value, i.e., a measure of the probability that the user will select a particular link or other screen object. After integration, the resultant pattern indicates in what degree each of the elaborated objects is related to the elaborated goal. Information scent is often defined narrowly as just semantic similarity, but here five factors combine to form information scent.

2.2.5 *Learning Mechanism of CoLiDeS*

CoLiDeS can learn in the same way as LICAI+ (Kitajima, Soto, & Polson, 1998), an extension of LICAI and a model of recall of occasionally used action sequences. CoLiDeS and LICAI store in memory episodes of successful instruction-following or correct step-discovery, recalling these episodes on ensuing occasions of encountering the same task to be performed on the same initial screen state. This process is analogous to recalling successfully solved word problem episodes when meeting the same problem again, and the resulting model of recall processes resembles models of text recall (Wolfe & Kintsch, 1998). LICAI+ adds to LICAI the processes involved in encoding and successfully retrieving encodings of correct actions. LICAI+ assumes that successful performance of occasionally performed tasks involves a mixture of recall of episodes of correct actions, and problem solving if recall fails. The model is related to Ross' (1984) and Rickard's (1997) models of skill acquisition. Kitajima et al. (1998) showed that LICAI+ predicts – and that data confirm LICAI+ predictions – that tasks that violate the label-following strategy are hard to learn by exploration and hard to remember, even if the correct steps have been previously presented.

2.3 Future of CoLiDeS and C-I Cognitive Architecture

The most serious problem we face is lack of a community of C-I modelers and an infrastructure to support researchers or practitioners who want to build HCI models in the C-I cognitive architecture. There are presently only limited models and tools and no full-fledged model development support comparable to the support for building cognitive models in ACT-R and Soar. LSA is a useful method for approximating the propositional networks of C-I, but the LSA community also offers limited support (http://lsa.colorado.edu). Does that mean that ACT-R or Soar will ultimately absorb C-I to incorporate current advantages of C-I theory of reading comprehension?

CoLiDeS has not yet been developed into a full, running simulation model, but some of our previous C-I models of HCI were full running simulation models (Kitajima & Polson, 1995, 1997). Creating a running simulation of CoLiDeS is high on our agenda and can be accomplished by adding mechanisms from LICAI, LICAI+, and LICAI/BT (briefly described in Kitajima & Polson, 2002). LICAI can use reading to understand task instructions. LICAI and LICAI+ can generate and select user goals. LICAI+ can learn action sequences from successful episodes, storing them as a cluster of propositions about a goal, elements of the context, and results of action. LICAI/BT can backtrack when a selection set (e.g., items on a pulldown menu) does not contain any objects similar enough to the current goal. LICAI originally used propositional representations, like C-I originally did, but Kintsch (1998) later began substituting LSA as an engineering approximation for propositional representations in C-I. CoLiDeS also uses LSA, and we could retrofit older members of our C-I family of models (LICAI, LICAI+, LICAI/BT) with LSA.

A cognitive architecture needs knowledge about how to use the interface and must be able to communicate with the interface. Embodied models (e.g., ACT-R/PM and ACT-R 5.0) can serve as simulated users. They must be able to interact with a representation of the same display that is presented to users/subjects and to interpret those displays and act on them in the same way that users/subjects do. Anderson (2002) explains that ACT-R traditionally focused on higher-level processes that could be broken down to the level of unit tasks that execute in about 10 seconds, but the newer features in the perceptual-motor interface extend ACT-R to deal with processes on a different level of magnitude – 50-millisecond slices. In contrast, C-I consistently started with the textbase and bypassed developing a theory of the low-level processes skilled readers use for visual recognition of letters and words. A single C-I cycle for sentence comprehension takes at least one second and usually a minimum of 2-3 seconds.

Thus, at present C-I cannot deal with the perceptual processes included in embodied models. C-I acknowledges that the skilled reader had to deal with orthography to get the textbase, but C-I starts with the textbase. Analogously, CoLiDeS acknowledges that visuospatial recognition ability is a mixture of top-down and bottom-up processes, but CoLiDeS starts with a webpage layout and text that is the equivalent of the textbase in C-I. CoLiDeS needs information about such conventional webpage widgets as the site search engine, site logo, and the representation of site information architecture in the left-hand navigation column. Our models (Kitajima & Polson, 1995, 1997) worked out interactions with pull-down menus, dialog boxes, etc., and all of this can generalize to any other screen display. Action planning mechanisms are powerful enough to interact with all the widgets on a webpage, and clicking links on a webpage is trivial by comparison with actions done in a Macintosh application. The roadblock is the lack of a formal theory of how users parse the webpage – lack of a well-developed way to represent a webpage that is equivalent to the way that Kitajima & Polson (1995) represented elements of Macintosh interface. Current

CoLiDeS research is filling the gap by investigating how webpage layouts direct attention. This research will extend the original definition of CoLiDeS (Kitajima, Blackmon, & Polson, 2000) in an important way by modeling the parsing process within the C-I architecture. Ultimately we must build a webpage parser that automatically parses the webpage the way human beings parse the webpage, representing a webpage layout in terms that CoLiDeS can use, including subregions, headings and subheadings, and links nested under the headings. There is some hope of doing this because the Web source code is open code (in contrast, most interfaces are closed systems), and so is the code for the browser and the rendering engine that produces the actual webpage layout.

Because it is difficult to master a complex cognitive architecture like C-I, and the cognitive models based on it (e.g., LICAI, CoLiDeS), our emphasis to date has been on developing CWW, an engineering approximation of the CoLiDeS cognitive model. CWW is simpler than CoLiDeS, usable by designers/practitioners, useful to them, and highly automated. We have created CWW tutorials that employ concrete examples and will continue to update these tutorials, make them available for download from our AutoCWW server, and add online help to server webpages.

3 Using LSA to Simulate Background Knowledge of Users

C-I architectures assume that perceptions of semantic similarity are dependent on successful text comprehension, that comprehension is dependent on having the necessary background knowledge, and that comprehension processes can fail due to inadequate background knowledge. We have consistently found strong evidence that inadequate background knowledge of a topic seriously impairs web navigation and contributes to high rates of task failure. Thus, we use the LSA cosine as a measure of semantic similarity, but it is only one of the LSA measures that we use. Furthermore, although LSA (Landauer & Dumais, 1997) is a very important component of CoLiDeS and CWW, CoLiDeS and CWW employ LSA under the guidance of the C-I cognitive architecture and its theory of comprehension processes. The guidance of C-I cognitive architecture has six crucial implications:

1. Because background knowledge is essential for comprehension, CoLiDeS starts by selecting an LSA semantic space that accurately represents the background knowledge of a particular user group in one language and culture (in two languages/cultures, if the user group is bilingual/bicultural).
2. Inadequate background knowledge impedes comprehension, and one feature of CoLiDeS that distinguishes it from SNIF-ACT is that CWW identifies topics and terms that will be unfamiliar to the user and given little or no attention. For example, the link Psychology has an adequate term vector length (0.97) for college-level general reading knowledge but drops to an unfamiliar level of 0.18 for the 9th-grade semantic space. Whenever a link text is unfamiliar, CWW and CoLiDeS assume that users are unlikely to select the link, even if LSA finds a strong semantic similarity between the user goal and the link.
3. Novel or low frequency words (words that occur infrequently in the corpus used to construct a semantic space) impair comprehension of link labels and webpage content. People probably ignore words and single-word link labels they have never seen (zero-frequency word). Analogously, LSA ignores words that have a zero frequency in the corpus, although this can precipitate the need to translate for LSA any recently coined word that real people know but LSA does not. Because people have only a vague sense of the meaning for words with very low frequencies, CoLiDeS flags words that have a frequency of 15 or less as words whose meaning must be inferred partially or completely from the surrounding context.
4. During comprehension a reader elaborates the text in relation to his or her background knowledge. Accordingly, CoLiDeS and CWW elaborate the raw text of each link in the selected semantic space by adding to the raw link text all the terms that have a minimum cosine of 0.50 with the raw text and a minimum word frequency of 50. For example, elaborating the raw link "Psychology" adds the terms "psychology psychologists psychologist sociology behavior anthropology psychological." Elaborated link labels generally produce more accurate estimates of semantic similarity (LSA cosine values).
5. Readers scan headings and subheadings to grasp the top-level organization or general structure of the text. While scanning, a webpage user can also locate a subregion that is above some threshold of similarity to the user's current goal, ignoring subregions that are not sufficiently familiar. To build a representation of each subregion, CoLiDeS and CWW first elaborates the raw heading text (like it elaborates each raw link text) and then it groups together the elaborated raw heading text and all the elaborated link texts for the links that are nested in the subregion. The LSA vector for the whole unit is an approximation of human understanding of the subregion and its heading, such as Geography or Social Science.
6. Different semantic spaces, representing different levels of education and/or different cultural backgrounds, produce different LSA cosine values as well as different LSA term vector lengths and word frequencies.

CoLiDeS differs from SNIF-ACT in its ability to accurately represent the background knowledge of a diversity of user groups by constructing and selecting an appropriate LSA semantic space. Thus, another future direction for CoLiDeS (see Section 2.3) will be to expand the variety of LSA semantic spaces and to update current LSA semantic spaces to reflect new knowledge acquired in recent years. In a rapidly changing world where information is increasingly accessed via the Internet, people are endlessly learning by exploration and seeking just-in-time information to accomplish their goals. Background knowledge differs widely among people from different cultural backgrounds, education levels, and areas of expertise, but these diverse groups are all using the same websites and the same Internet. CoLiDeS focuses on the conditions necessary to support pure forward search and to identify and repair impediments to pure forward search, so CoLiDeS must accurately predict how successfully diverse users can navigate websites to accomplish particular goals. The accumulated evidence about diverse users can, in turn, contribute to further developing the C-I cognitive architecture, expanding the range of evidence to a far broader sample of human beings. In addition, if personas (Cooper, 1999) prove highly useful to HCI designers, an appropriate semantic space could be attached to each persona – linking semantic spaces with concrete images of particular users who might use any particular website. During ethnography for developing a persona, the usability expert could collect texts representing what a persona has read and use them to create a specialized semantic space.

4 Engineering Approximation of CoLiDeS for Website Usability Evaluation

CWW (Blackmon et al., 2002, 2003, 2005) is a simplified version of CoLiDeS, an engineering approximation designed to be a pragmatic website usability evaluation method. Usability analysts must find and repair usability problems rapidly, and CWW must be designed so that analysts can rapidly learn to use CWW accurately. CWW must be highly automated and must substitute objective measures for subjective human judgments, particularly in cases where the human users have very different background knowledge than the usability analysts.

- CWW now has automated algorithms to identify usability problems. CWW simulates the two pairs of C-I cycles, first parsing the webpage and selecting a subregion to attend to, and then selecting a link from the attended-to subregion. CWW identifies a competing heading problem if attention is strongly drawn to an incorrect subregion(s) rather than a correct one that contains a link that leads to accomplishing the goal.
- CWW now accurately predicts the mean total clicks to accomplish a particular goal on a webpage and can, thus, predict task difficulty and determine what tasks have problems serious enough to warrant repair.
- CWW suggests repairs that allow users to find items by clicking the links they are most apt to click. Users vary in the links clicked to reach the goal, so to create successful repairs CWW must encompass enough high-probability links to make it possible for almost all users to reach the goal within two or three clicks.
- The CWW prediction formula is most accurate for the most serious problems, and the success rate for repairs is best for the most serious problems. Thus CWW enables usability experts to target the most serious problems for repair, the problems that are most worth the time and effort required to repair them.

5 Summary of Experiments Testing CWW Predictions

We have conducted experiments that demonstrate the psychological validity and reliability of CWW problem identification, the accuracy of the CWW measure of problem severity, and the high success rate of CWW-guided repairs of problems (Blackmon et al., 2002, 2003, 2005). In these studies we used predicted mean total clicks as the measure of problem severity, but in the present context it is more useful to focus on factors that foster pure forward search. Two groups of experimental participants alternated between tasks that CWW predicted to be non-problems and tasks that CWW predicted to be serious problems, completing a total of 64 tasks searching for an article on a simulated online encyclopedia. Both groups, one with 53 experimental participants, the other with 60, were composed of undergraduates who completed the tasks for partial course credit. The main webpage for each of the 64 tasks displayed 93 links nested under nine headings, and only one of the 93 links was a correct link that led to the target article, the same link that leads to the article in the online encyclopedia being simulated in the experiment.

Table 1: Minimum-path solvers and observed clicks for non-problem and problem tasks

Task Type	Number of Tasks	Mean percentage of minimum-path solvers. (standard deviation)	Mean observed total clicks (standard deviation)
Non-problem	34	59.54%. (20.39)	2.2 (0.89)
Problem	30	17.38%. (18.98)	5.6 (1.91)

Table 1 shows the results. For each task we computed the percentage of participants who completed the task in the minimum number of clicks: a mean percentage of 60% minimum-path solvers for the 34 tasks that CWW predicted would be non-problem tasks compared to a mean percentage of only 17% minimum-path solvers for the 30 tasks that CWW predicted would pose serious problems. The difference was statistically significant, $F(1, 62) = 72.7$, $p < .0001$. Table 1 also shows the mean observed total clicks for the same tasks and subjects: 2.2 clicks for non-problem tasks vs. 5.6 clicks for problem tasks, a statistically significant difference, $F(1, 62) = 89.3$, $p < .0001$. Thus, virtually everyone finds non-problem tasks in 1-3 clicks, approaching an asymptote of pure forward search. Multiple regression analyses for this dataset and for a larger set of 164 tasks (Blackmon et al., 2005) demonstrated three factors responsible for explaining variance in task difficulty (indexed by mean observed total clicks): presence of an unfamiliar correct link, presence of a weak-scent correct link, and number of competing links nested under competing headings. These are precisely the factors tracked by CoLiDeS. Research on CWW (especially Blackmon et al., 2005) offers solid empirical verification for CoLiDeS based on controlled experiments, collectively offering a consistent pattern of evidence from hundreds of tasks, each task completed by a large sample of subjects that ensures a stable mean for each task.

6 Conclusions

Due its C-I foundation, CoLiDeS has unique features that enable it to model reading comprehension, the core cognitive process underlying webpage navigation, and to accurately predict usability problems that impair navigation. Both CoLiDeS and SNIF-ACT deal with information scent, but only CoLiDeS manipulates the effects of background knowledge and identifies problems due to inadequate background knowledge – unfamiliar topics and low frequency words. We have repeatedly demonstrated (Blackmon et al., 2002, 2003, 2005) that inadequate background knowledge (unfamiliar problems) impedes pure forward search and cause serious problems for users. Inadequate background knowledge impairs successful navigation even for users with college-level general reading knowledge, but gaps in background knowledge expand for users with less adequate background knowledge, e.g., readers represented by the 6th-grade semantic space instead of the college-level space. CoLiDeS can simulate qualitative as well as quantitative differences among semantic spaces, e.g., differences in cultural backgrounds.

Another key feature of CoLiDeS is its two pairs of C-I cycles that first parse the webpage and select a subregion as the focus of attention. In terms of Information Foraging theory, CoLiDeS selects one patch within the webpage and ignores the rest. If the correct link is in the attended-to patch, as occurs in non-problem tasks, the selective attention fosters pure forward search and efficiently curtails wasting effort on attending to unpromising information patches. As we have shown (Blackmon et al., 2002, 2003, 2005), when subregions (information patches) within a webpage have high scent for a user goal but fail to contain a link leading to the goal, they act as competing headings and tend to cause high rates of task failure. SNIF-ACT currently processes the links on a webpage sequentially, so it has no mechanism for detecting situations where users are drawn to the wrong information patch within the webpage.

To ensure a promising future of C-I architecture in HCI, the bottom line is empirical verification. Though CoLiDeS is not yet a full running simulation like SNIF-ACT is, CoLiDeS has already compiled a large amount of empirical evidence to demonstrate its psychological validity and reliability. CoLiDeS and other models based on C-I have thus carved a promising niche in HCI. At this juncture, nothing would teach us more about the consequences of cognitive architectures than head-to-head testing of a model derived from one cognitive architecture compared to a competing model derived from another cognitive architecture, e.g., SNIF-ACT from ACT-R compared to CoLiDeS from C-I..

References

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science, 26,* 85–112.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

Blackmon, M.H., Kitajima, M., & Polson, P.G. (2005). Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. *Proc. of Conference on Human Factors in Computing Systems*, *CHI 2005*.

Blackmon, M.H., Kitajima, M., & Polson, P. G. (2003). Repairing usability problems identified by the Cognitive Walkthrough for the Web. *Proc. of Conference on Human Factors in Computing Systems*, CHI 2003, 497–504.

Blackmon, M.H., Polson, P.G., Kitajima, M., & Lewis, C. (2002). Cognitive Walkthrough for the Web. *Proc. of Conference on Human Factors in Computing Systems*, *CHI 2002,* 463–470.

Chi, E. H., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions and the Web. *Proc. of Conference on Human Factors in Computing Systems*, *CHI 2001*, 490–497.

Chi, E.H., Rosien, A., Supattanasiri, 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. *Proc. of Conference on Human Factors in Computing Systems*, *CHI 2003*, 505–512.

Cooper, A. (1999). *The inmates are running the asylum: Why high tech products drive us crazy and how to restore the sanity.* Indianapolis, IN: Sams Publishing.

Engelbeck, G.E. (1986). *Exceptions to generalizations: Implications for formal models of human-computer interaction.* Unpublished Masters Thesis, Department of Psychology, University of Colorado, Boulder.

Franzke, M. (1994). *Exploration, acquisition, and retention of skill with display-based systems.* Unpublished Ph.D. Dissertation, Department of Psychology, University of Colorado, Boulder.

Franzke, M. (1995). Turning research into practice: Characteristics of display-based interaction. *Proc. of Conference on Human Factors in Computing Systems*, *CHI'95*, 421–428.

Furnas, G.W. (1997). Effective view navigation. *Proc. of Conference on Human Factors in Computing Systems*, *CHI'97*, 367–74.

Kitajima, M., & Polson, P.G. (2002). LICAI/BT: A comprehension-based model of item selection with backtracking capability. *International Conference on Soft Computing and Intelligent Systems*.

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! (Proceedings of HCI 2000,* 357–373*)*. Springer.

Kitajima, M., Soto, R., & Polson, P.G. (1998). LICAI+: A comprehension-based model of the recall of action sequences. *Proc. of 2nd European Conference on Cognitive Modeling*, 82–89.

Kitajima, M., & Polson, P. (1997). A comprehension-based model of exploration. *Human-Computer Interaction*, 12, 4, 345–389.

Kitajima, M., & Polson, P. (1995). A comprehension-based model of correct performance and errors in skilled, display-based human-computer interaction. *International Journal of Human-Computer Studies*, 43, 65–99.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, U.K. & New York: Cambridge University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review*, 95, 163–182.

Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge, *Psychological Review* 104(2), 211–40.

Mannes, S.M. & Kintsch, W. (1991). Routine computing tasks: Planning as understanding. *Cognitive Science*, 15, 305–342.

Pirolli, P. (2004). The use of proximal information scent to forage for distal content on the World Wide Web. In A. Kirlik (Ed.), *Working with technology in mind: Brunswikian resources for cognitive science and engineering*.

Pirolli, P., Fu, W.-T., Reeder, R., & Card, S. K. (2002). A user-tracing architecture for modeling interaction with the World Wide Web. *Advanced Visual Interfaces (AVI'02)*, Trento, Italy, May 22–24.

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

Pirolli, P. L., & Fu, W-T. (2003). SNIF-ACT: a model of information foraging on the World Wide Web. *Proc. of 9th International Conference on User Modeling* (*Lecture Notes in Artificial Intelligence 2702),* 45–54.

Polson, P. G., & Lewis, C. (1990). Theory-based design for easily learned interfaces. *Human–Computer Interaction*, 5, 191–220.

Rickard, T.C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, 126, 288–311.

Ross, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371–416.

Wolfe, M.B.W., & Kintsch, W. (1998). Memory for text: Background knowledge and text-base influences. Poster presented at the meeting of the Psychonomic Society, Dallas, TX.