



# NAVIGATION-DEPENDENT WEB VIEWS : ACTIVATING AND COMPLEMENTING WEB LINKS FOR A PERSONALIZED WEB

Kiyomitsu, Hidenari

---

(Citation)

Kobe University Economic Review, 45:67-78

(Issue Date)

1999

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81000922>

(URL)

<https://hdl.handle.net/20.500.14094/81000922>



# NAVIGATION-DEPENDENT WEB VIEWS: ACTIVATING AND COMPLEMENTING WEB LINKS FOR A PERSONALIZED WEB

By HIDENARI KIYOMITSU

In this paper, we propose an idea for changing the view of web-pages and their link structures based on a user's link navigation history and predefined *semantic units* of web pages. The major objective of the proposed method is the reflection of a web page author's intention about his/her web data and its linking structures. In our *navigation-dependent web views*, the view of each web page including its hyperlinks is changed according to which navigation path the user took to reach the web page. If the user obeys a navigation path that is intended by the author, the system will provide a necessary and sufficient view of the web page including all the predefined hyperlinks. Otherwise, our mechanism adds information on unvisited pages to the user's view and hides some hyperlinks. In order to realize the proposed navigation-dependent web views, first we will introduce the notions of *semantic units* and *situation-dependent links*. Intuitively, a *semantic unit* is predefined by its author, and it consists of a connected graph of web pages, where each edge denotes a prerequisite relationship between pages as well as a hyperlink. That is, in order to comprehend the content of a semantic unit, users are strongly recommended to enter from a predefined entry page and to navigate predefined links in the specified order. When a user does not navigate the links according to the author's intention, some hyperlinks are automatically hidden so that he may not be able to go to further pages. These types of hyperlink we called *situation-dependent link*. Next, we will briefly discuss our implementation of the navigation-dependent web views. Finally, we will briefly discuss the formal aspects of the *situation-dependent links*. We will formalize the syntax and semantics of semantic units of pages. Then, we will define the equivalence relationships between two semantically meaningful paths. The equivalence relationship is useful for finding redundant paths defined by authors.

## 1. Introduction

Hypertext is a powerful tool for expressing relationships between documents by hyperlinks with ease. A vast number of documents is stored on the Internet sites, and they are connected through hyperlinks. Hyperlinks are very flexible, and users can freely connect their documents with other documents by establishing hyperlinks.

It is also typical for an author to create his/her web pages assuming that their users will follow an implicit navigation path. For example, when an author creates a page, say A, followed by a page B, then the author assumes that readers will navigate from page A to page B following the predefined hyperlink. However, WWW users can create hyperlinks from their favorite pages to page B directly, in which cases the author's intention is disregarded, and page A is not always read by users.

In order to prevent such undesirable navigation by users, the adoption of the following

- **Semantic Units**

A web document author can specify a collection of related web pages as a *semantic unit*. Within a semantic unit, the readers of the web pages can obtain the full

information if they obey the predefined navigation paths.

- **Situation-Dependent Links**

According to a user's actual navigation path, which may not be the same as the author intends, the system dynamically controls the visibility of hyperlinks. If a reader does not obey the predefined navigation paths, the system automatically also shows the contents of unvisited pages to him, and/or the system makes some hyperlinks invalid.

In Section 2, we describe the notion of *semantic units*. Section 3 describes the notion of *situation dependent links*. In Section 4, we show our prototype implementation and examples. In Section 5, we show some formalization of semantic units and situation dependent links. Section 6 is a discussion of several problems which we should further consider. Section 7 gives concluding remarks.

## 2. Semantic Units

Suppose that a web document author is creating a document with three web pages, say A, B, and C. In order to understand the whole contents of his document, he may wish the readers of the web pages to read pages A, B, and C in that order. In order to reflect the author's intention, he will establish a hyperlink from A to B, and a hyperlink from B to C. In this sense, the aim of these hyperlinks is not only to navigate the three web pages, but also to enforce a kind of *prerequisite* relationship among web pages. That is, page A should be read before page B, and page B should be read before page C in order to comprehend the whole contents of the pages.

Figure 1 (i) shows simple web pages and their hyperlinks. There is no restriction on how to navigate pages A, B, and C. Readers can directly read page B by specifying the URL of page B. In the same manner, as shown in Figure 1 (ii), readers can freely access either A, B, or C in an arbitrary order, and some readers may read only page B or C independently from page A.

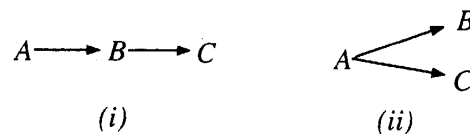


FIGURE 1. Usual Hyperlinks

On the other hand, if a hyperlink from page A to page B conveys the prerequisite relationship between page A and B as well as the usual navigation relationship, it is necessary to represent the semantics in a different manner. In this paper, we will denote by a box a collection of hyperlinks having the prerequisite relationship semantics. Each hyperlink surrounded by a box denotes a prerequisite relationship between two web pages. The collection of nodes and hyperlinks surrounded by a box is called a *semantic unit*. A semantic unit is a meaningful unit of web pages such that the comprehension of the web pages requires the reader to navigate all the web pages by navigating predefined links.

For example, the box shown in Figure 2 (i) indicates that pages A, B, and C consti-

tute a semantic unit of information, where the understanding of page B requires the reading of page A, and the understanding of page C needs the comprehension of page B. In the same manner, the box in Figure 2 indicates that pages A, B and C constitutes a semantic unit of web pages and that readers are recommended to first read page A, and then to read pages B and C (in either order) to understand the whole contents of the semantic unit.

More formally, a semantic unit is a collection of hyperlinks such that those hyperlinks convey both the usual navigation relationships and prerequisite relationships. For example, a semantic unit for Figure 2 (i) is a collection (A, B), (B, C) of hyperlinks. A semantic unit for Figure 2 (ii) is a collection (A, B), (A, C) of hyperlinks. Note that even if we had another hyperlink (A, D) in figure 2 (ii), the hyperlink (A, D) would not be always contained in the semantic unit. Also, the hyperlinks appearing in a semantic unit must comprise a single connected graph.

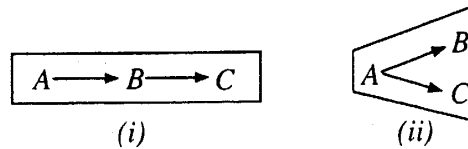


FIGURE 2. Semantic Units

### 3. Situation Dependent Links

Web users (readers) can access any web pages if they know the corresponding URLs in advance. So, some users can directly access a web page in a semantic unit that is NOT an entry point to the semantic unit. That is, direct access to some pages in a semantic unit does not match the author's intention.

In this section, we discuss our concept of *situation-dependent links*, which controls the visibility of hyperlinks according to the reader's navigation path, and complement information on unvisited page for readers so that they may not lose any information.

#### 3.1 The Basic Concept of Situation-Dependent Links

Suppose that an author defines a semantic unit  $\{(A, B), (B, C), (C, D)\}$  as shown in Figure 3. But, other web users can access to the page C via an undesirable link from a page F, which was established by somebody else. In this case, the author hopes to represent the meaning of the hypertext document consisting of pages A, B, C and D in order. Thus browsing only C and D from F via this undesirable hyperlink is not what the author intended. But, it is not desirable to reject any access to page C except that from B, because we are afraid this rejection may reduce the flexibility of hypertext fatally.

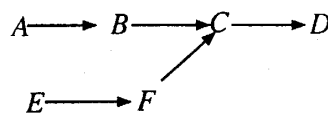


FIGURE 3. Undesirable Hyperlink

Our concept of a situation-dependent link is illustrated in Figure 4. Our basic idea is to control the visibility of hyperlinks based on a user's navigation history. That is, a hyperlink from a node within a semantic unit is made *visible* (and viable) if and only if a reader already navigates all the prerequisite nodes within the same semantic unit. This hyperlink, whose visibility is controlled by the reader's visiting history, is called a situation-dependent link.

As shown in Figure 4, we have a single semantic unit  $\{(A, B), (B, C), (C, D)\}$ . Here, each oval is used to denote a user's navigation history. Also, each ellipse is a hyperlink visualized by a user navigation history. If a user has already navigated pages A, B, and C, the hyperlink (C, D) becomes visible since the user already visited all the prerequisite pages. On the other hand, suppose that another user has already visited pages E, F, and C. Then, the hyperlink (C, D) from page C is NOT visible because the prerequisite pages A and B have not been visited by the user.

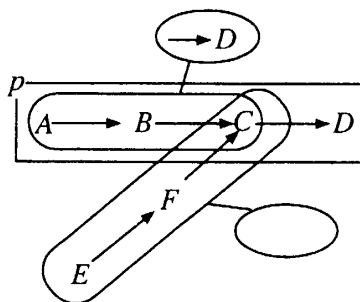


FIGURE 4. Situation Dependent Link

In this example, the semantic unit  $p = \{(A, B), (B, C), (C, D)\}$  is a path from page A to page D and the order of  $p$  is a sequence of pages by which the author intends a user to browse. Now let a user  $x$ 's navigation path be  $navigation(x)$ , and if he has browsed pages A, B and C, in this order then let his navigation path be denoted by  $navigation(x) = ABC$ . Here ABC is an abbreviation of a path (A, B), (B, C). Also, let another user  $y$ 's navigation path be  $navigation(y) = EFC$ . Then, the condition for a user to be able to reach page D is to have the navigation path ABC in  $navigation(user)$ . User  $x$  can reach page D because his navigation history  $navigation(x)$  is ABC, but user  $y$  can not reach page D. In other words, a path to page D becomes valid by browsing from the first page in a semantic unit including page D to the page previous to page D. This has the semantics; "understanding the meaning of page D requires the understanding (comprehending) of pages A, B and C". Since the user  $x$ 's navigation history has all the elements to page D, the situation-dependent link (C, D) to page D becomes viable dynamically.

### 3.2 Complementation of Information Using Situation Dependent Links

Our goal is avoiding the misunderstanding of the meaning of hypertext documents consisting of multiple pages. The concept of the semantic unit and the situation-dependent link is not only a matter of hiding hyperlinks from any users accessing via undesirable links. It is also possible to provide complementary information based on semantic units

with those users. That is, the basic idea is automatically to complement the information on unvisited pages to users who are conducting undesirable navigations.

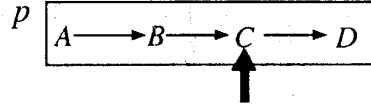


FIGURE 5. Complementation of Unvisited Pages

Figure 5 shows an access to an intermediate page C within a semantic unit  $p = \{(A, B), (B, C), (C, D)\}$  against the author's intention via a bold arrow hyperlink. By using the previous situation-dependent links, the immediate access to page C does not make the link (C, D) visible because of the existence of the semantic unit  $p$ . This results in the situation that the user cannot go to page D.

On the other hand, our notion of *information complementation* is to require the user to read pages A and B when he or she directly enters page C. That is, our system tries to guide the user to visit the prerequisite pages by showing page C together with the URLs of pages A and B. This is an example of *positive* usage of our semantic units.

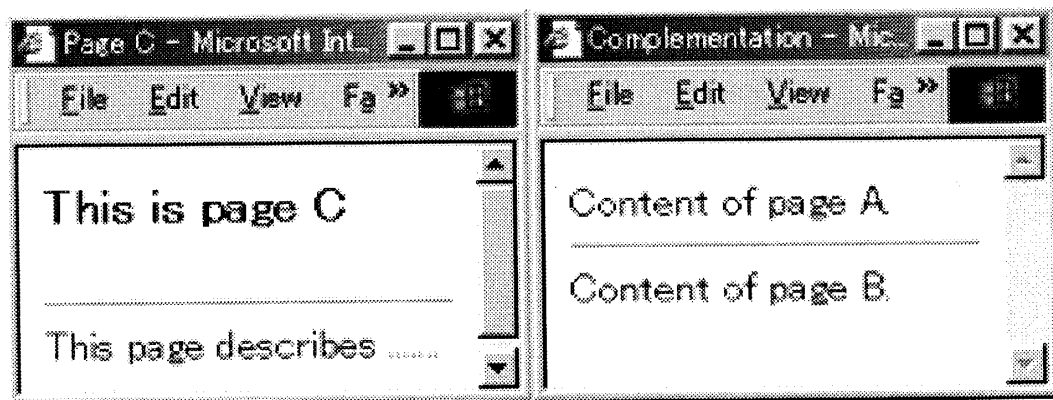


FIGURE 6. Complementing Skipped Pages

For example, as shown in Figure 6, when a user visits page C directly, the complementation mechanism may show the skipped (unvisited) pages A and B or only the URLs of pages A and B in an additional window while a user is browsing page C. Only after reading the contents of pages A and B, does the situation-dependent link (C, D) of the semantic unit  $p$  become visible and it guides the user to page D.

The following is a general procedure for finding which information should be complemented when a user accesses an intermediate page  $v$  within a semantic unit:

1. Find a collection of pages immediately previous to page  $v$ . (Suppose that we have a collection  $\{u_1, u_2, \dots, u_m\}$  of pages such that each  $u_i$  is a page immediately previous to page  $v$ .) If there are no pages immediately previous to page  $v$ , then no information is complemented.
2. For each  $u_i$  in  $\{u_1, u_2, \dots, u_m\}$ , find *minimal* semantic units including page  $u_i$ .
3. For each minimal semantic unit found at Step 3, all the pages from each entry point up to page  $u_i$  are shown to the user who is accessing page  $v$ .

As shown in Figure 7, assume that we have the following four semantic units:

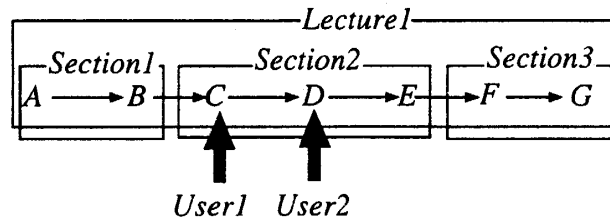


FIGURE 7. Example of Lecture

- *Lecture 1* = {(A, B), (B, C), (C, D), (D, E), (E, F), (F, G)}
- *Section 1* = {(A, B)}
- *Section 2* = {(C, D), (D, E)}
- *Section 3* = {(F, G)}

This is an example in which several semantic units are arranged in a hierarchical manner. The meanings of the semantic units are obvious. *Lecture 1* is a series of web pages A, B, C, D, E, F, G and it is divided into three sections. Suppose that an immediate access by *User 1* to page C in semantic unit *Section 2* occurs. The page immediately previous to page C is B. Then, the minimal semantic unit that contains page B is *Section 1*. Since the entrypoint to *Section 1* is page A, pages A and B are shown to *User 1* complementarily.

Next, assume that *User 2* accesses page D. Then, the page immediately previous to page D is page C. The minimal semantic unit that contains page C is *Section 2* = {(C, D), (D, E)}. The entrypoint to *Section 2* is page C, and so, only page C is shown to *User 2* as complementary information.

#### 4. Implementation and Examples

The implementation of the semantic units, situation-dependent links, and information complementation function in a typical web environment. Recently, hypertext languages have been given some facilities for specifying active links that are embedded procedures, and many user agents can pursue their facilities. But it is still difficult to obtain user's navigation histories from clients. Because there is no hypertext language (including XML) to store a user's navigation history. Netscape Navigator can store a user navigation history in a file "`~/ .netscape / history.db`" on UNIX platform, but lack of permission to use this file prevents anyone from getting a user's history. Therefore it is difficult to store a user's history on the client side.

Figure 8 illustrates a system architecture for implementation. Our solution for implementing the situation-dependent links and information complementation uses Server Side Includes (SSI). Any page in a semantic unit has a procedure for storing and checking a user's visiting history and for suggesting the next page or one or more complements. In our implementation, a suggestion of a situation dependent link is represented as a hyperlink to the next page or a set of hyper links to prerequisite pages. But, the method of visualization of this information depends on the author's objective and implementation. Because of the use of HTML as a hypertext language in this implementation, situation-dependent links are embedded in a visited page as in-line links.

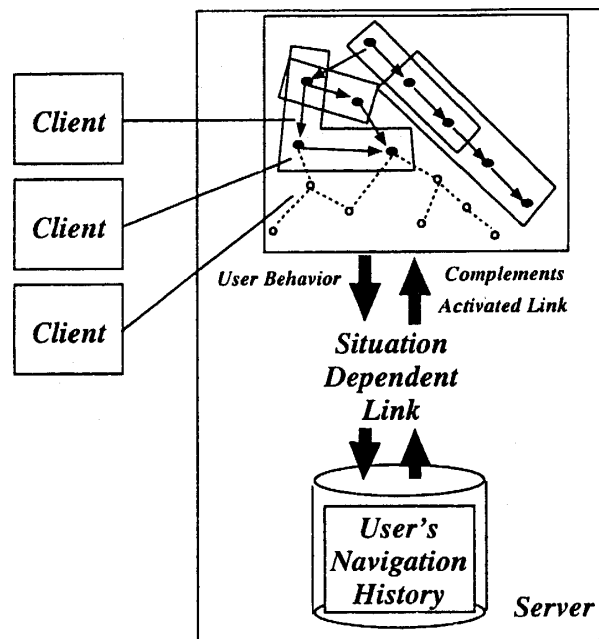


FIGURE 8. An Example of System Structure

On the server side, a situation-dependent link proceeds according to a user's link navigation. When a user accesses a page in semantic units, the server receives environment variables such as ENV. An element of ENV indicates an IP address as a symbol 'REMOTE\_ADDR'. Our implementation uses an IP address for identifying the user. This is why the e-mail address of the user or user ID is not always sent in issuing the "http get" from a client. Therefore user visiting history data is a file enumerating IP addresses. First, the system evaluates an IP address sent from a client if it exists in an IP addresses file of the visited page as the behavior of a client. Then, the system returns one or more hyperlinks to the next pages or complements.

This is an atomic methodology, but it is useful for PC client environments. Figure 9 illustrates an example of undesirable access to page C in the semantic unit *Lecture 1*. Complements for understanding this page are suggested above its contents in a window. So, the situation-dependent link of the semantic unit *Section 2* found a minimal semantic unit *Section 1* required to understand page C, and provided these complements as two inline links to page A and page B. When a user navigates the hyperlink to page A by using the complements in page C, his navigation history is stored, and the browsing situation changes. His visiting history data at page C is stored in a history file of page C. Our system uses this file when any pages relevant to this page as prerequisites in semantic units are accessed. In this situation, his visiting history data is C, A. Therefore, this situation dependent link suggests a hyperlink navigation to the user.

Next, the user navigates page B, and the situation for browsing page C changes. When he visits page C again, the situation dependent link visualizes a hyperlink to page D as a nested page of C. Because his navigation history has changed (at this time he has visited pages C, A, B and C), the prerequisites for his visit to page D are now fully satisfied. This situation is illustrated in Figure 12. The difference between Figure 9 and

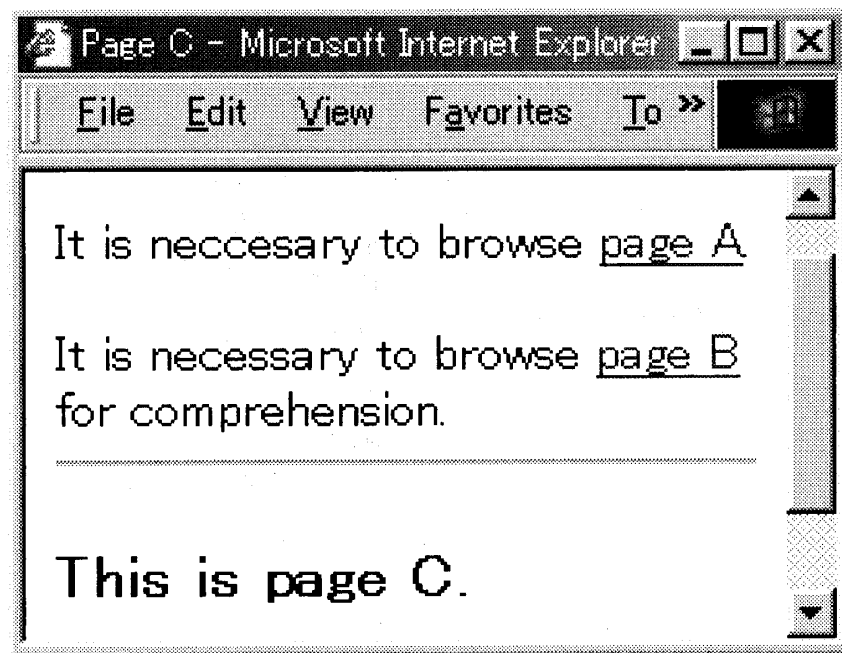


FIGURE 9. Undesirable Access from Page C

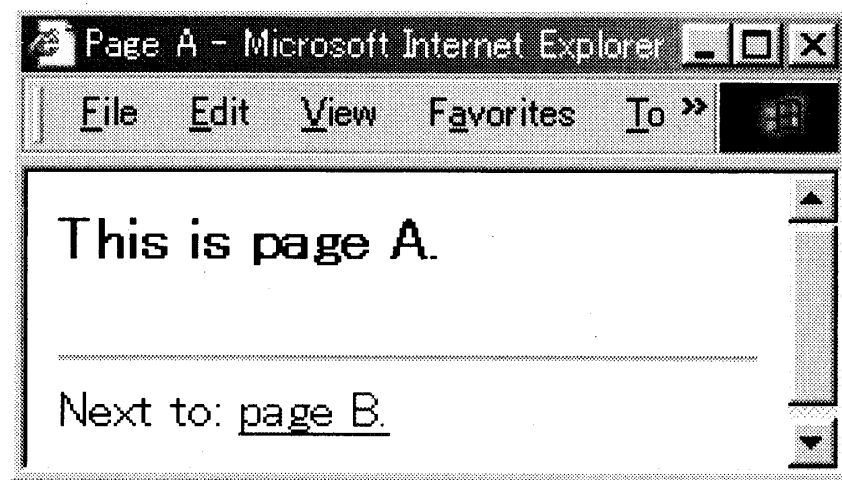


FIGURE 10. User's Navigation to Page A

Figure 12 indicates that the situation dependent links reflect the author's intention for user navigation.

### 5. Formalization: Navigation-Admissible Paths

In this section, we will briefly discuss the equivalence relationship between two sets of semantic units. Hereafter, for simplicity, we restrict every semantic unit to a path of edges (note that in general, a semantic unit is a general graph).

The *navigation-admissible paths* of a semantic unit  $p = \{(v_1, v_2), (v_2, \dots), \dots$ ,

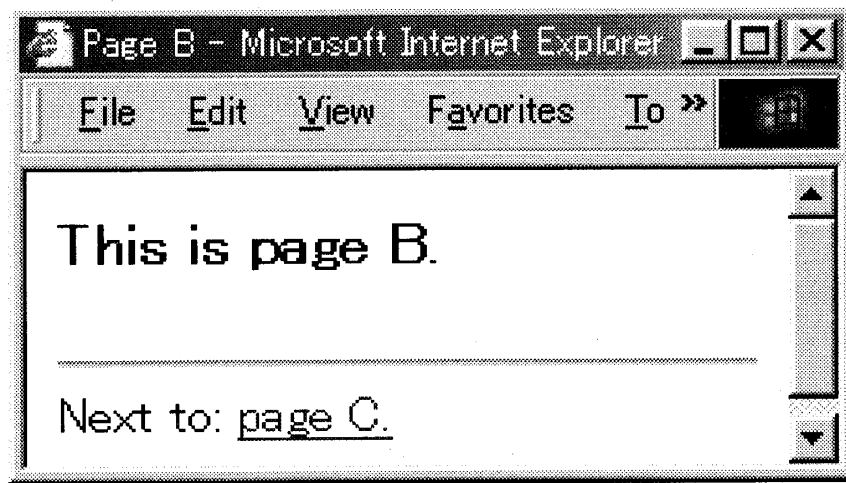


FIGURE 11. A View of Page B in this Example

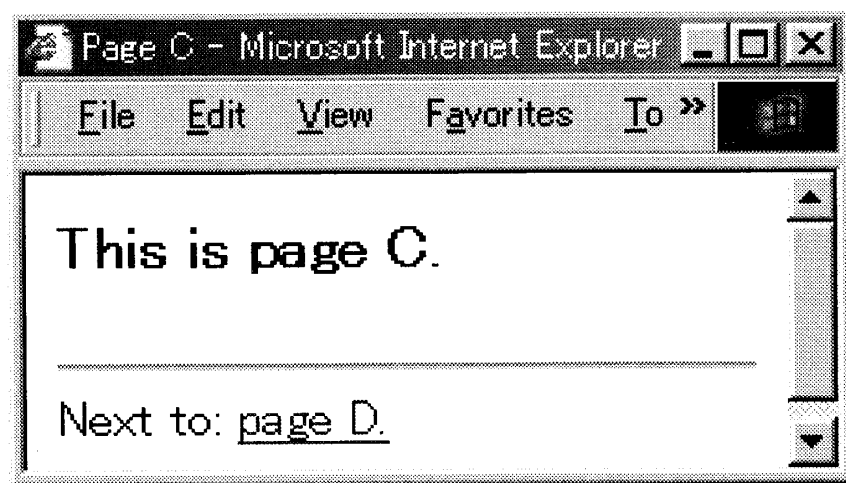


FIGURE 12. User's Navigation to the Page C

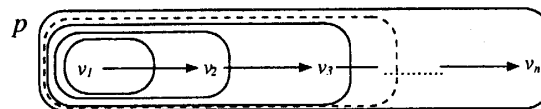


FIGURE 13. Navigation-Admissible Path

$(v_{n-1}, v_n)$  are denoted by  $Path(p)$ , and are defined as sets of all the desirable paths within  $p$ . That is, for  $p$ ,  $Path(p)$  is a set  $\{(v_1, v_2), (v_1, v_2), (v_2, v_3), \dots, (v_1, v_2), \dots, (v_{n-1}, v_n)\}$  (see Figure 13). For given semantic units  $p$  and  $q$ ,  $p$  is said to be equivalent to  $q$  if and only if  $Path(p) = Path(q)$  holds. More generally, for given sets  $P$  and  $Q$  of semantic units,  $P$  is said to be equivalent to  $Q$  if and only if all the desirable paths in  $P$  are equal to those in  $Q$ .

As shown in Figure 14, let's consider the case that a semantic unit  $p_1$  includes another semantic unit  $p_2$ . Also, let  $P$  and  $Q$  be  $P = \{p_1, p_2\}$  and  $Q = \{p_1\}$ , respectively. Then, obviously, the navigation admissible path  $Path(P)$  of  $P$  is equal to  $Path(p_1) \cup$

$Path(p_2)$ . Also,  $Path(P) = Path(Q)$  holds. Therefore, the two semantic units  $P$  and  $Q$  are said to be equivalent because both of these have the same navigation admissible paths.

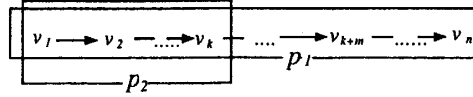


FIGURE 14. Relationship between  $p_1$  and  $p_2$

On the other hand, we should be careful to consider a case when a semantic unit overlaps with other semantic units, as shown in Figure 15. Here, let  $Q = \{p_1, p_2\}$  be a set of a semantic unit. Also, let  $P = (v_1, v_2), \dots, (v_{n-1}, v_n)$  be a semantic unit. In this case, clearly,  $Path(\{P\})$  is NOT equal to  $Path(Q)$  since  $Path(\{P\})$  does not contain a path  $(v_k, v_{k+1})$ .

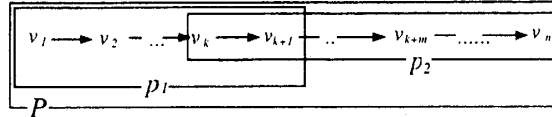


FIGURE 15. Overlapping  $p_1$  and  $p_2$

## 6. Discussion

We have discussed a user's navigation over semantic units and the complementation of unvisited page information by using situation-dependent links. But some problems of the treatment of information units exist on WWW.

A semantic unit is a directed acyclic graph and comprises a single connected graph. Therefore its nodes can be graphs. We consider some situation-dependent links from a semantic unit to another semantic unit, from a page to a semantic unit and from a semantic unit to a page or pages as shown in Figure 16.

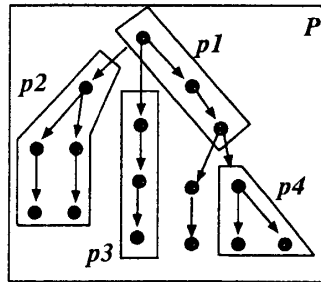


FIGURE 16. Nested Graph Structure

A semantic unit  $P$  is a directed acyclic graph with some subordinates. One  $P$ 's situation-dependent link to the first page of  $p_2$  is from a semantic unit  $p_1$ . The semantics of this link are: "in order to understand the first page of  $p_2$ , it is necessary to comprehend the whole semantic unit  $p_1$ ". The other  $P$ 's situation-dependent link to the semantic

unit  $p_4$  is from the last page of  $p_1$ . The semantics of this linkage are: “in order to understand the meanings of  $p_4$ , it is necessary to comprehend the last page of  $p_1$ ”. An example of Figure 17 is a hierarchical version of Figure 7. In this example, a semantic unit *Lecture 1* does not consist of pages A, B, C, D, E, F but of *Section 1*, *Section 2* and *Section 3*.

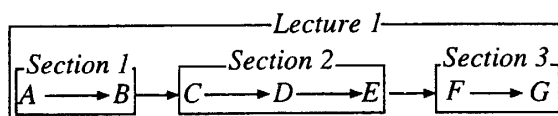


FIGURE 17. Hierarchical Version of Lecture 1

The document structure of semantic unit *Lecture 1* is illustrated below. **SEQ** is a mnemonic symbol of the sequential relationship between these sections and pages in a specific logical structure on Open Document Architecture (ODA/ODIF) [7].

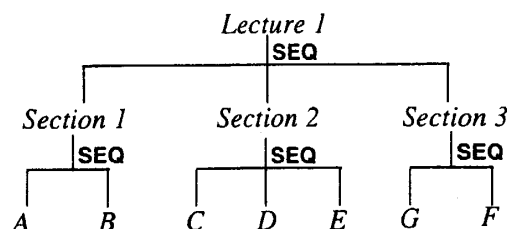


FIGURE 18. Document Structure of Figure 17

## 7. Conclusions

In this paper, we have proposed the idea of *navigation-dependent web views*, which can change the view of web pages and their link structures based on the notions of *semantic units*, *situation-dependent links*, and *information complementation*. Also, we have briefly described our prototype implementation and a formalization of these ideas. It should be noted that our ideas can be implemented without any sacrifice of the freedom of the web access mechanism. In our *navigation-dependent web views*, the view of each web page including its hyperlinks is changed according to which navigation path the user took to reach the web page. If the user obeys a navigation path that is intended by the author, the system will provide a necessary and sufficient view of the web page including all the predefined hyperlinks. Otherwise, our mechanism adds information on unvisited pages to the user's view and hides some hyperlinks. We have shown the ways to add information and to hide hyperlinks based on the notion of semantic units. Finally, under the restriction that every semantic unit is a path, we briefly discussed the formal aspects of semantic units such as the equivalence relationships between two semantic units.

Recently, similar ideas were proposed by Stotts and Cabarrus [2] and by Réty [3]. Stotts and Cabarrus regard the web as an automaton, and hyperlink navigation as controlled by a user's state (which is determined by the user's navigation history). Réty also adopts the same approach, and the visibility (or the navigability) of hyperlinks are

controlled by previously-visited pages. In particular, constraints concerned with the link visibility are represented by PROLOG-like logic programs. These two approaches are the same as our work in spirit, but, neither of them introduced the notion of *semantic units*. Furthermore, in comparison with these works that are automata-based or logic-based, our approach is a graph-theoretic approach.

One more direction of related research is concerned with how to *find* semantic units in a web graph. Tajima et al. [4] proposed a way to discover querying units in graph-structured data such as the web, netnews, and E-mails. Hatano et al. [5] and Li et al. [6] showed a way to find a semantic information unit, that is a subgraph of web, for a given query. Clearly, their approaches are different from ours because our work assumes that semantic units are determined a priori by their authors. But, our notions of situation dependent links and information complementation can be applied to those semantic information units found by their approaches.

Finally, the following issues are to be further studied:

- General formalization of semantic units and related issues.
- Development of more-effective presentations of complemented information.
- Implementation of navigation-dependent web views for an XML environment.

### Acknowledgements

Thanks are due to Dr. Katsumi Tanaka, Dr. Keishi Tajima, Dr. Kazutoshi Sumiya, and other members of the Tanaka Lab. of Kobe University for their useful discussions. This research is partly supported by Research for the Future Program of the Japan Society for the Promotion of Science under the Project: Research on Advanced Multimedia Contents Processing.

### REFERENCES

- [1] Hidenari Kiyomitsu and Katsumi Tanaka (1999), "Navigation-Dependent Web-Views: Defining and Controlling Semantic Units of Web Pages," *IEEE Proc. of ICSC '99*, LNCS 1749, pp. 421-426.
- [2] P. David Stotts, Cyrano Ruiz Cabarrus (1998), "Hyperdocuments as Automata: Verification on Trace-Based Browsing Properties by Model Checking," *ACM Transaction of Information Systems*, Vol. 16, No. 1, pp. 1-30.
- [3] Jean-Hugues Réty (1999), "Structure Analysis for Hypertext with Conditional Linkage," *Proc. of Hypertext '99*, pp. 135-136.
- [4] Keishi Tajima, Yoshiaki Mizuuchi, Masatsugu Kitagawa, and Katsumi Tanaka (1998), "Cut As a Querying Unit for WWW, Netnews, and E-mail," *Proc. of Hypertext '98*, pp. 235-244.
- [5] Kenji Hatano, Ryouichi Sano, Yiwei Dan and Katsumi Tanaka (1999), "An Interactive Classification of Web Documents by Self-Organization Maps and Search Engines," *Proc. of DASFAA '99*, pp. 19-25.
- [6] Wen-Syan Li and Yi-Leh Wu (1998), "Query Relaxation by Structure for Web Document Retrieval with Progressive Processing," *Proc. of Advanced Database Symposium '98*, pp. 19-25.
- [7] ISO/IEC 8613-2 (1994), "Open Document Architecture and Interchange Format (ODA/ODIF): Document Structures".