METHOD OF ANNOTATING DISPLAYS AND AN ANNOTATION MODULE

This Sheet Has Been Replaced

Field of the Invention

The present invention relates to a method of annotating displays and an annotation module, which are particularly, but not exclusively, useful for annotating Internet documents.

Background of the Invention

The worldwide computer network known as the Internet is primarily based on the "client-server" model of information exchange. In this "distributed computing" environment, a server (host), which is normally a powerful computer or group of computers, behaves as a single computer and services the requests of a large number of smaller computers, or clients, which connect to it.

The Internet supports a large variety of information transfer protocols. Of these, the Hypertext Transfer Protocol (http) which supports the World Wide Web (the "web") is probably the most prominent. An important feature of the web is the ability to connect one file, or web page, to many other pages using "hypertext" links. A link appears either as an underlined or highlighted portion of text, or simply as part of an image object in a document. When a viewer of a web page moves the cursor over a hyperlink and clicks, the link is executed and the linked file retrieved, and that file need not be located on the same server as the original file.

A client computer typically retrieves documents on the web using a browser, such as Netscape Navigator™ or Microsoft Internet Explorer™, which utilises a HTML interpreter to execute HTML instructions to display a page.

THIS SHEET HAS BEEN REPLACED

This Sheet Has Been Replaced



The number of web pages accessible by a client computer is enormous and constantly growing. However, a characteristic of the web is that all web pages are created, maintained, and delivered by the hosts. Clients which connect to the hosts simply display the web pages by interpreting the embedded HTML commands using the browsers. Users on the Internet are therefore effectively readers of information presented by editors of the web pages on the hosts, who are often referred to as "webmasters".

As the web is accessible by millions of users with a great diversity in language, culture, religion, interests, literacy and training, a webmaster simply cannot cater to every user's needs. A typical webmaster will therefore choose to satisfy the largest number of target users possible, and this usually means presenting the content using American English. In some cases, a webmaster serving a small group of target users may choose to use language specific to the group. For example, a web page intended for medical doctors' consumption may contain a lot of medical terms which render the document incomprehensible for users without appropriate training. Users on the web can therefore be frustrated by his or her inability to understand the content of the web pages. This defeats the original purpose of the web being a huge information resource.

Summary of the Invention

In accordance with the present invention there is provided a method of annotating displays, including:

requesting display data;

receiving and processing said display data for display items which require annotation;

generating a display, using said display data, with said display items including respective annotations in said display; and

accessing and displaying item data for one of said display items on selection of the respective annotation for said one of said display items.

The present invention also provides an annotation module stored on a computer readable medium, including an interception module for receiving and processing requested display data for display items which require annotations, and adding annotation data to said display data to cause generation of a display, using said display data, with said display items including respective annotations in said display.

THIS SHEET HAS BEEN REPLACED

The present invention also provides a computer apparatus for annotating displays, including:

means for receiving and processing display data for display items which require annotation;

means for requesting said display data and generating a display, using said display data, with said display items including respective annotations in said display; and

means for accessing and displaying item data for one of said display items on selection of the respective annotation for said one of said display items.

Brief Description of the Drawings

Preferred embodiments of the present invention are hereinafter described, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a block diagram of a preferred embodiment of a computer system including an annotation module; and

Figure 2 is a block diagram of components of the annotation module.

Detailed Description of the Preferred Embodiments

An annotation module 4, as shown in the Figures, operates in tandem with a conventional document-retrieval facility, such as a web browser 45, by altering requested document data , such as web pages, in real-time after they have been retrieved from the hosts, but before being displayed to the users. The module 4 annotates the documents with links which can access additional explanation of selected words. The module 4 scans the retrieved documents and identifies words which are likely to be incomprehensible to the user, and annotates these words, typically with inconspicuous bullets next to these words. The annotated documents are then passed on to a display device, such as the web browser 45, for display to the user. While reading the document, the user can choose to click on any bullet if additional information about the associated word is required. The explanation can contain a translation, word sense, or example of usage or any other appropriate reference which is displayed in a separate window.

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

The annotation module 4 monitors a user's reaction to the annotations and retains use data for effectively learning the user's preference. The use data is used subsequently to determine how a word should be annotated in subsequent documents. User preference is determined from the use data by the cumulative number of times an annotation is retrieved or ignored, or by an explicit user statement. The module 4 also allows dictionaries to be updated or replaced.

A computer system 2, as shown in Figure 1, which includes the annotation module 4, can operate as a network client and may be a personal computer running WINDOWS™. The system 2 includes a bidirectional bus 20, over which all system components communicate, at least one mass storage device (such as a hard disk or optical storage unit) 22, and a main system memory 24. Operation of the system 2 is directed by a central-processing unit (CPU) 26. A conventional communication platform 30, which includes suitable network interface capability and transmission hardware, facilitates connection to and data transfer through a computer network 31, such as the Internet, over a telecommunication link 33. The user interacts with the system using a keyboard 35 and a position-sensing device (e.g. a mouse) 37. The output of either device can be used to designate information or select particular areas of a screen display 39 to direct functions to be performed by the system 2.

The main memory 24 includes a group of executable software modules that control the operation of CPU 26 and its interaction with the other hardware components. The modules include the annotation module 4 and the web browser 45. An operating system (not shown), such as WINDOWS 95™, directs the execution of low level, basic system functions such as memory allocation, file management and operation of mass storage device 22, multitasking operations, input/output and basic graphics functions for output on screen display 39. The user's primary interactions with the system occur using the web browser 45, which contains functionality for locating and fetching, via the network 31, data, such as web pages, each identified by a Universal Resource Locator (URL), temporarily storing and displaying these, executing hyperlinks contained in web pages and selected by the user, and generally interpreting web page information.

THIS SHEET HAS BEEN REPLACED

This Sheet Has Been Replaced



This Sheet Has Been Replaced



The annotation module 4 includes a control module 48 which interacts with the web browser 45, a web interface 50, which interacts with the communications platform 30, and a dictionary module 52 which includes dictionary entries for various terms, which include words or phrases. The control module 48 accesses web items from the web interface 50, which communicates with the communication platform 30 and stores retrieved web page data in the manner of a web browser 45. The web interface 50 provides the same interface to the communications platform 30 which the web browser 45 would normally provide, for example an interface which works with Winsock ™ that may be part of the communications platform 30.

The control module 48, as shown in Figure 2, includes an interception module 104, a usage analysis module 102 and a term preference module 106. The interception module 104 annotates data or web items of a web page using the term preference module 106 and the dictionary module 52. The interception module 104 processes the web page data stored by the web interface 50 in real-time as it is passed to the web browser 45, so that a user of the browser 45 is unable to notice any latency introduced by the interception module 104 in generation of the web page display. The web page data is scanned by the module 104 for all text items which may require annotation. Using the accessed text items, the interception module 104 consults the term preference module 106 to determine if the user prefers to have an item or term annotated. The term preference module 106 designates terms by weights or values which indicate the level of user preference. Terms with preference values lower than a threshold are considered "non-preferred" to indicate that a user does not require them to be annotated. An accessed term which is not marked "non-preferred" and which has a corresponding entry in the dictionary module 52 is selected for annotation. Selection is performed in real-time using either automata, hashing tables or a quick tree-search when processing the web page data. Dictionary entries accessible by the dictionary module 52 include respective hyperlinks to data relating to a term, which normally comprises explanatory text or a translation. If a term is selected, the interception module 104 inserts a display bullet with the accessed appropriate hyperlink in the processed web page data so that the bullet will appear close to the annotated web item. The interception module 104 then passes the annotated page to the web browser 45 for display.

Once a page has been annotated, a user is able to access item data corresponding to a web item by using the position-sensing device 37 to select the corresponding bullet and associated hyperlink. On selection of the link, the request for the item data passes via the

THIS SHEET HAS BEEN REPLACED

This Sheet Has Been Replaced



usage module to access a HTML file which may be maintained by the dictionary module 52 or which can be generated by a program stored on the system 2. The file contains the requested item data, which normally would be stored on the system 2. The file is accessed via the browser 45 in the same manner as for any other hyperlinked file.

The usage analysis module 102 observes how a user interacts with the inserted annotations or bullets. Based on the observations, the usage analysis module 102 sends messages to the term preference module 106 to update the preference values which indicate the level of user preference. Actions monitored by the usage analysis module 102 include:

1. The cumulative number of times the user has clicked on bullets associated with a term.
2. The cumulative number of times the user has ignored bullets associated with a term.
3. Whether the user has saved a reference to the term, e.g. by adding the term using the usage analysis module 102 to a host term list maintained by the term preference module 106. Such action represents an explicit indication of interest in the term.

Raw data associated with the above is stored within the term preference module 106 for further processing. The primary purpose of term preference module 106 is to decide whether a user is interested to have a term annotated. The annotation module 4 can be provided with means for the user to manually reset the term preference module 106, clear all previous preferences, and change the threshold value which determines whether a term is considered "non-preferred". This enables the user to signal a complete change of interest.

The dictionary module 52 enables the user to manually add, modify and delete any dictionary entries, or substitute a completely different dictionary. This capability enables the user to have different terms annotated, or have the same term annotated with different types of dictionary entries.

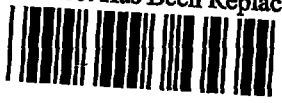
All of the components and modules of the annotation module 4 can be implemented using the Java programming language, so that they can either be installed locally, or loaded and executed as required.

THIS SHEET HAS BEEN REPLACED

The annotation module 4 is not limited to use on the Internet. The architecture described above can, for example, be used directly with local area networks of computers communicating via, for example, the Ethernet protocol. In a local area network, the computers can implement TCP/IP over the low level Ethernet hardware management routines to create an intranet, or can instead (or in addition) be tied into the Internet as a node, via, for example, a telephone hookup to an external host computer serving as a commercial Internet service provider. Alternatively, the system can be used with other forms of document-viewing facility (whether these involve a computer network or a single machine) by replacing web interface 50 with an appropriate retrieval system, so as to alter retrieved data in real-time before being displayed to a user.

Many modifications will be apparent to those skilled in the art without departing from the scope of the present invention as described herein.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED



This Sheet Has Been Replaced



CLAIMS

1. A method of annotating displays, including:

requesting display data;

receiving and processing said display data for display items which require annotation;

generating a display, using said display data, with said display items including respective annotations in said display; and

accessing and displaying item data for one of said display items on selection of the respective annotation for said one of said display items.

2. A method as claimed in claim 1, wherein said annotations comprise links to respective item data.

3. A method as claimed in claim 1, wherein said processing of said display data and adding said annotations thereto occurs in real-time

4. A method as claimed in claim 2, wherein the method is executed on a machine connected to a communications network.

5. A method as claimed in claim 4, wherein said display data is received from said network and said item data is accessed from said machine.

6. A method as claimed in claim 1, including:

monitoring selection of said annotations;

storing preference data; and

determining display items which require annotations on the basis of said preference data.

7. A method as claimed in claim 6, wherein said preference data includes a number of respective times annotations have been selected, and a list of said display items maintained by a user.

8. A method as claimed in claim 1, wherein said display items include terms in

THIS SHEET HAS BEEN REPLACED



documents and said item data includes dictionary entries for said terms.

9. A method as claimed in claim 4 or 5, wherein said network is the Internet.

10. A method as claimed in claim 9, wherein said display data is web page data.

11. An annotation module stored on a computer readable medium, including an interception module for receiving and processing requested display data for display items which require annotations, and adding annotation data to said display data to cause generation of a display, using said display data, with said display items including respective annotations in said display.

12. An annotation module as claimed in claim 10, wherein said annotation data includes links to respective item data for said display items, such that selecting the respective annotation causes display of respective item data.

13. An annotation module as claimed in claim 11, wherein said interception module processes said display data and adds said annotation data in real-time.

14. An annotation module as claimed in claim 12, wherein the computer readable medium is part of a machine connected to a communications network.

15. An annotation module as claimed in claim 14, wherein said display data is received from said network and said item data is accessed from said machine.

16. An annotation module as claimed in claim 15, wherein said network is the Internet.

17. An annotation module as claimed in claim 16, wherein said display data is web page data.

18. An annotation module as claimed in claim 12, including a preference module storing preference data which is used by said interception module to determine display items which require annotations.

This Sheet Has Been Replaced



THIS SHEET HAS BEEN REPLACED

19. An annotation module as claimed in claim 18, including a usage module for monitoring selection of said annotations and causing storage of said preference data in said preference module.

20. An annotation module as claimed in claim 19, wherein said preference data includes a number of respective times annotations have been selected, and a list of said display items maintained by a user.

21. An annotation module as claimed in claim 12, including an item module for maintaining and providing access to said item data.

22. An annotation module as claimed in claim 21, wherein said display items include terms in documents and said item data includes dictionary entries for said terms.

23. An annotation module as claimed in claim 11, including an interface module for receiving and storing said requested display data for said interception module.

24. A computer apparatus for annotating displays, including:

means for receiving and processing display data for display items which require annotation;

means for requesting said display data and generating a display, using said display data, with said display items including respective annotations in said display; and

means for accessing and displaying item data for one of said display items on selection of the respective annotation for said one of said display items.

25. A computer apparatus as claimed in claim 22, wherein said annotations comprise links to respective item data.

26. A computer apparatus as claimed in claim 24, wherein said processing means process said display data and adds said annotations thereto in real-time.

27. A computer apparatus as claimed in claim 25, wherein said apparatus is connected to a communications network.

28. A computer apparatus as claimed in claim 27, wherein said display data is received

THIS SHEET HAS BEEN REPLACED

from said network and said item data is stored on said apparatus.

29. A computer apparatus as claimed in claim 25, including:
means for monitoring selection of said annotations; and
means for storing preference data for use in determining said display items which require annotations.

30. A computer apparatus as claimed in claim 29, wherein said preference data includes a number of respective times annotations have been selected, and a list of said display items maintained by a user.

31. A computer apparatus as claimed in claim 30, wherein said display items include terms in documents and said item data includes dictionary entries for said terms.

32. A computer apparatus as claimed in claim 27 or 28, wherein said network is the Internet.

33. A computer apparatus as claimed in claim 32, wherein said display data is web page data.

This Sheet Has Been Replaced

* C L M S *

THIS SHEET HAS BEEN REPLACED