

Web Annotator

Dale Reed, Sam John
Computer Science Department
University of Illinois at Chicago
Chicago, IL 60607-7053
reed@uic.edu, sjohn@cs.uic.edu

Abstract

The World Wide Web is increasingly becoming an integrated extension of users' computing environments, with content indexed and retrieved through Web browsers. Web browsers are increasingly being used as computer science curriculum delivery mechanism, for both books delivered as local content on CD ROMs as well as server-based material.

Traditional computer science curriculum has often been presented through static printed media. What has been printed ahead of time in books or handouts can not be changed. Any changes would have to be attached externally by way of scribbled notes or explanations on the board. The Web gives us the technological affordances to change that, both for students and teachers.

The Web Annotator is an Internet Explorer plug-in that allows users to annotate browser-based course material. It allows teachers to cooperatively develop course content, as well as provide supplemental annotations to course web pages. Material on *any* web page can be added, deleted, or reformatted.

The original page remains unchanged, however the user's version of the page is filtered, adding, subtracting and changing web page elements, giving an enhanced view. A demonstration version of the Web Annotator can be downloaded from:

<http://logos.cs.uic.edu/Annotator>

Categories & Subject Descriptors

K.3.2 [Computers & Education]: Computer & Information Science Education - Computer Science Education.

General Terms

Documentation, Design, Human Factors.

Keywords

Web-based Curriculum, Annotation, Collaborative Design, Browser Plugin.

Permission to make digital or hand copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, require prior specific permission and/or a fee.
SIGCSE'03, February 19-23, 2003, Reno, Nevada, USA.
Copyright 2003 ACM 1-58113-648-X/03/0002...\$5.00.

1 The Need

Web browsers such as Netscape and Internet Explorer are increasingly becoming common vehicles for computer science curriculum delivery. Back in 2000, one in three U.S. colleges offered an accredited degree online [10], with this figure doubtless being much higher today. Curriculum which has traditionally been in books is also being moved to an electronic format. For instance several U.S. dental schools have replaced traditional textbooks with a single DVD containing the curriculum for all four years of school [11]. In computer science some instructors eschew traditional texts in favor of a collection of on-line resources.

The ability to customize the view of on-line curriculum is important for both educators and students. Educators can use the Web Annotator to organize or annotate existing on-line curriculum from other courses or partner institutions, enhancing the examples to make them better situated for the students' programming ability.

Students can use the Web Annotator to annotate on-line material as well as to synthesize excerpts from different web pages. For instance a student in an introductory programming lab connected to the Internet during a lab session could make notes to themselves on course web pages directly rather than writing them down on a separate piece of paper. This close linking provides a meaningful context for the notes.

2 The Solution

The solution, quite simply, is to allow editing any web page, saving the differences from the original page as a filter. This allows uses such as:

1. A list of commands can be customized to fit a particular programming environment (adding new ones and deleting others).
2. Additional examples personalized to the students' context can be inserted into an existing set of examples.
3. Teachers can collaboratively develop web-based pages, sharing their annotations through a central server.

The first step is to download the browser plug-in which self-extracts and installs as an optional browser toolbar. Once installed, the components on the annotation toolbar will appear as shown in **Figure 1**.

Any web page can be modified by selecting the *Edit* button. This puts the document into edit mode, where changes – additions, deletions, and format modifications - can be made. Once all the

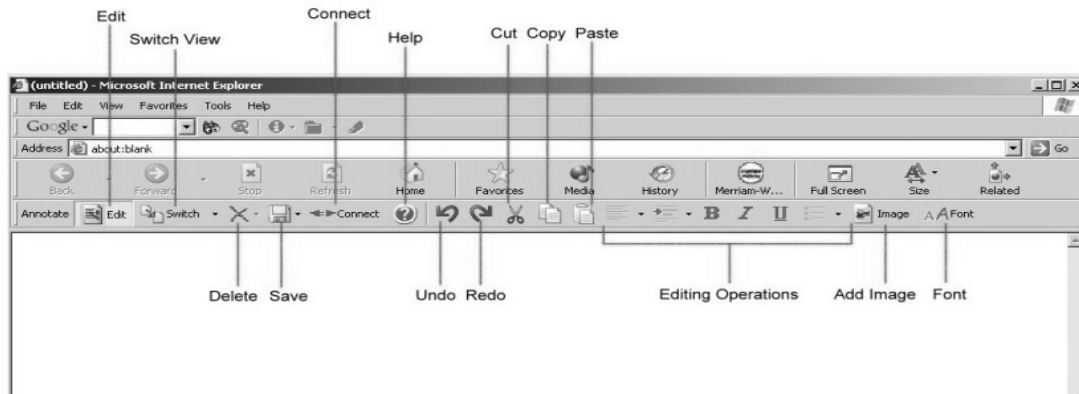


Figure 1: Web Annotator Toolbar

changes have been made, changes are saved by selecting the *Save* button. The *Switch View* button allows comparing the original and modified views, where the modified views can be stored locally (the default) or on a shared server. Using a shared server allows teachers to collaboratively design or comment upon a page, where selecting different views presented in the drop-down menu under the *Switch View* button allows a user to quickly compare multiple versions of a page.

Current annotations allow deleting and adding both text and images. Fonts can be changed, and justification and formatting can be applied. Following these buttons we find the buttons for undo/redo, cut, copy, paste, etc. The Web Annotator requires IE 5.5 or higher to run. The program is fully functional as a standalone product, as the ability to share filters through a server is not required. The server database uses *mySQL*, a free database product, as one of the design goals was to make this product low or no cost.

It is important to point out here that the page to be annotated need not be under the control of the user making the annotations. It can be any page from any curriculum source anywhere on the Internet or in a local file, as long as it comes up in your browser. The changes are saved as html text along with the corresponding URL. With filters enabled, subsequent visits to the original URL automatically apply the stored customizations.

3 Web Annotator in Action

Consider the example of customizing a curriculum page that is giving examples and explanations of the *if* statement in the C programming language. The appearance of the original page is shown in **Figure 2**, with the enhanced page with applied customizations shown in **Figure 3**.

In this example the user first enabled editing, then made the following changes:

- Numbered the “Common Problem” topics
- Highlighted the line with the error in bold red
- Added a note below explaining that the code will not generate a compiler error.
- Inserted a picture beside the note, to give it more weight

Note that the original page is not changed. Only the user’s view of it is changed.

4 Still to Come

Still to come is the ability to allow creation of a composite page, or collage made up of excerpts from multiple Internet sources. These have the sources automatically attributed. Existing web page building tools (e.g. Front Page, Dreamweaver) allow building a new page using discrete components of an existing page, such as a graphic image, but do not allow extracting a subset of text to be used as a component of the new page unless that text is copied. Figure 3 shows what this process would look like.

The user would first create the destination page. In this case it is the “Variations on the *if* statement” page shown in the middle of **Figure 4**. The user would then highlight areas of other web pages to be added, copying and pasting them into new page. Attribution (the [source] link) would automatically be added. The way this would be done is to create a composite view of the individual pages, each in their own frame. Everything would be filtered out except the sections the user highlighted for inclusion. Following the [source] link would load the original page.

5 Problems

There is the obvious potential for a copyright problem if users were to take other’s work and represent it as their own. Also some web page designers intend for their work to always appear in its entirety. These are ongoing issues that exist in cases where users outright copy sections of others’ pages and paste them into their own documents. There is a sense in which we are attempting to avoid this misrepresentation of others’ work by building in the “bibliographic entry” into each element that would be added to a page.

The Web Annotator also depends on the underlying content being fairly static, as it uses text/image positioning to know what to change. Drastic changes to the original underlying pages could render the filters ineffective.

The Web Annotator is currently not portable to other browsers or platforms (e.g. Netscape or Macintosh). A demonstration copy can be downloaded from:

<http://logos.cs.uic.edu/Annotator>

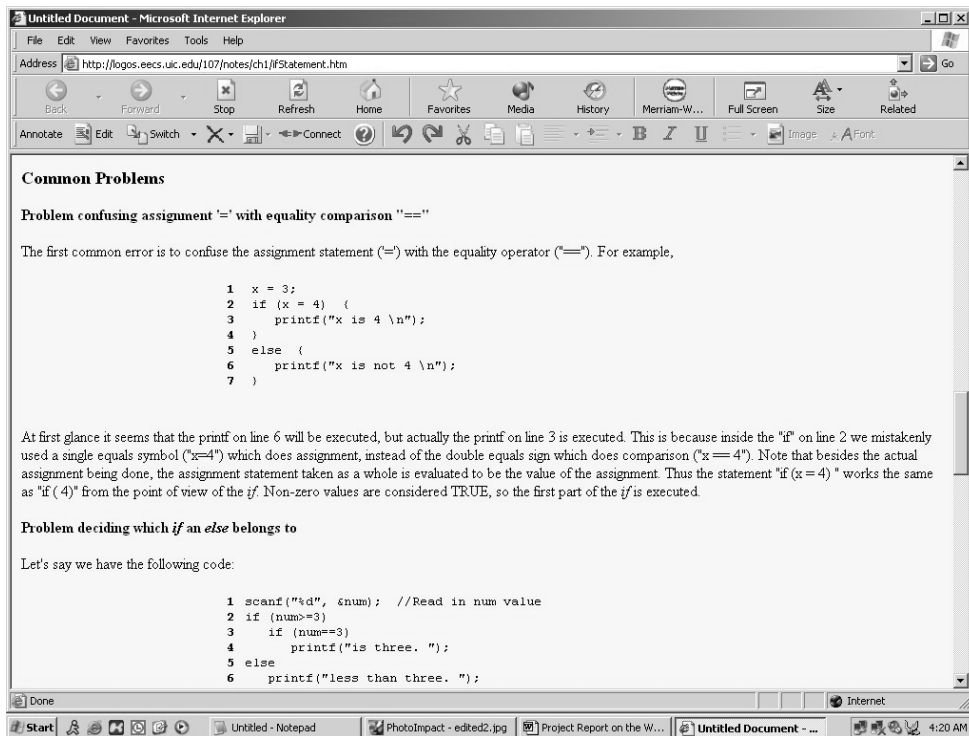
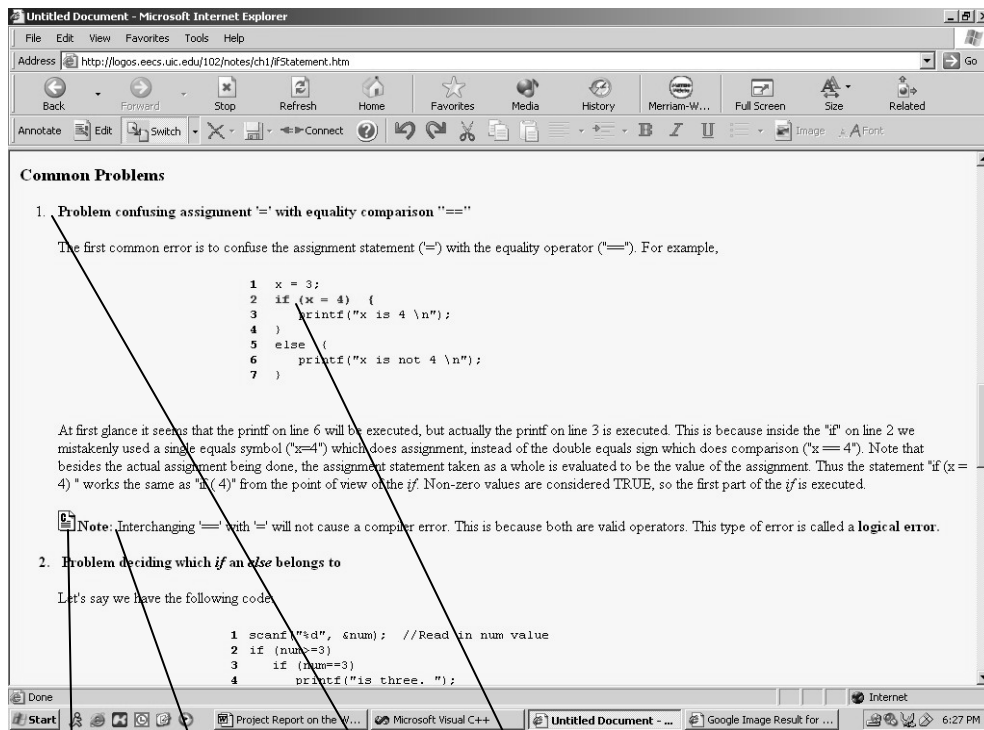


Figure 2: Original Page



Highlighted the line containing the error in bold red

Indented and numbered the common errors

Added a text note clarifying a concept

Added a picture indicating something to notice

Figure 3: Changed Page

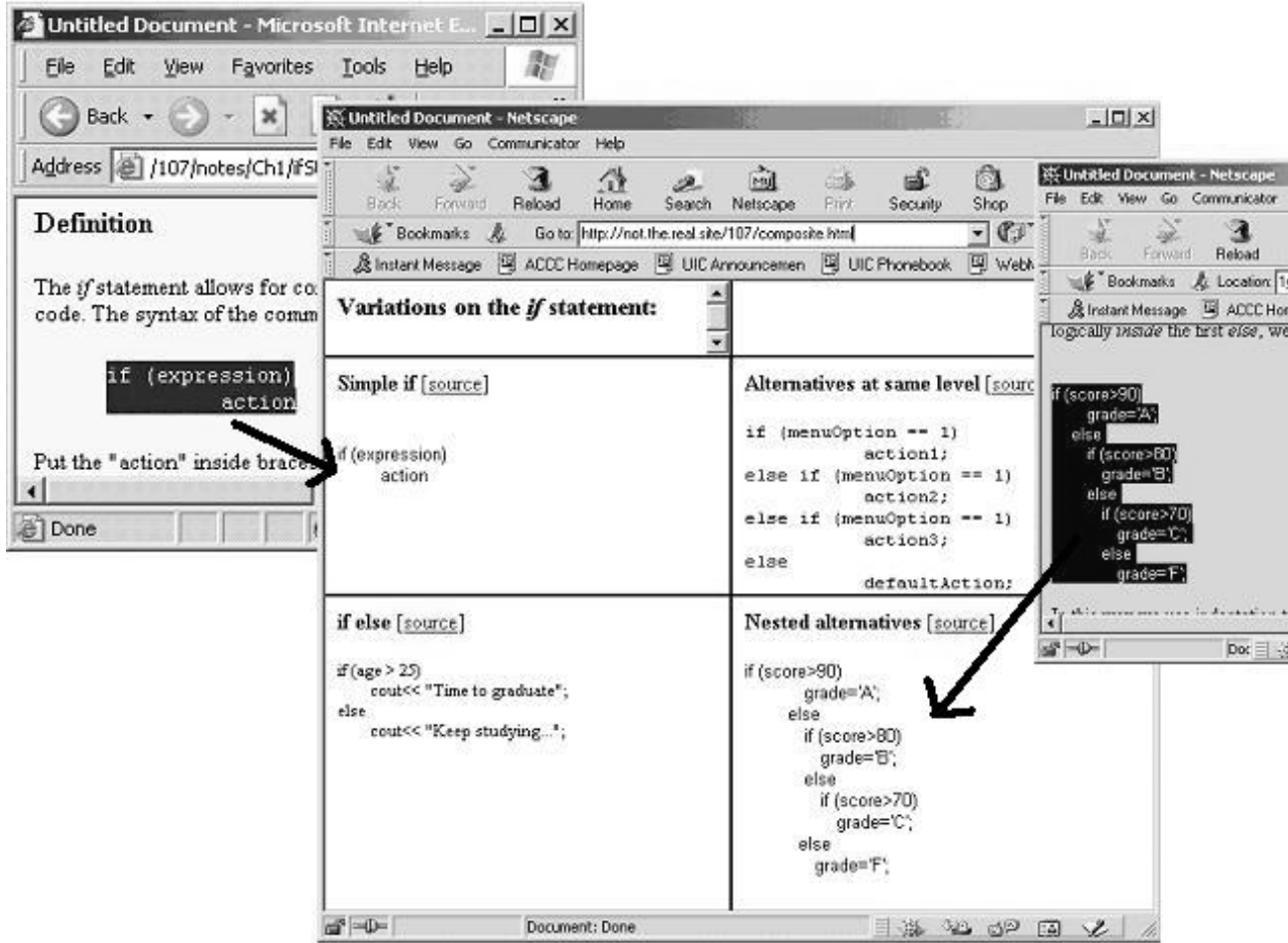


Figure 4: Composite “collage” page with automatic attribution.

6 Related Work

There are other filtering and annotating tools that have been developed, but none have integrated the ability to modify the user’s view of others’ web pages in the way that this work proposes. Additionally it is currently possible in designing a web page to refer to a graphic image using a link to the corresponding graphics file (“deep linking”), but it is not possible to do this with an arbitrary excerpt of text, as we propose to do.

Most annotating software to date is server based [2]. Notable among these was ThirdVoice [9], which allowed annotating any web page, though it had to be done through ThirdVoice’s server. As of this writing, ThirdVoice has discontinued their service. EZsite WebNotes [5] allows side-by-side annotation of a document, but that document must be under control of the user. The Interactive Paper Project (IPP) [3] allows commenting on a paragraph at a time on documents that are stored on the IPP server. Server based solutions offer the advantage of sharing annotations, but tend to be slow.

The star of the client-based web document annotators is iMarkup [7] which allows users to post notes, highlight text, and draw freehand on the surface of any web page. All information is stored on the users own machine. Software Post-It Notes [4] act as a customizable web page bookmark that lives on your desktop, while Hot Off the Web [6] allows capturing a web page and then annotating it for possible subsequent distribution.

The use of a proxy server to act as an intermediary between a client’s web browser and the HTML delivered by a server has also been variously discussed [1,8,12].

This work was supported in part by NSF’s PFSMETE program, NSF 2-5-28025.

References

- [1] Barrett, R. and Maglio, P.P. Intermediaries: new places for producing and manipulating Web content. Internet [September 4, 2002]. WWW: www7.scu.edu.au/programme/fullpapers/1895/com1895.htm
- [2] Hummes, J., Karsenty, A., and Merialdo, B. Active Annotations of Web Pages. Internet [March 30, 2000]. Available WWW: www.eurecom.fr/~hummes/docs/Web4Groups/w4g.html
- [3] Internet [September 4, 2002]. WWW: lrsdb.ed.uiuc.edu:591/ipp Forms based.
- [4] Internet [September 4, 2002]. WWW: www.3m.com/market/office/postit/com_prod/webnotes/index.html
- [5] Internet [September 4, 2002]. WWW: www.dougdean.com/EZsiteWebNotesSite.htm
Document must have a server-side include of a Microsoft IIS ASP component
- [6] Internet [March 30, 2000]. WWW: www.hotofftheweb.com
- [7] Internet [September 4, 2002]. WWW: www.imarkup.com IE only
- [8] Internet [September 4, 2002]. WWW: Muffin.doit.org has java-based proxy-server filtering programs
- [9] Internet [March 30, 2000]. WWW: www.ThirdVoice.com Runs on both Netscape and IE.
- [10] Los Angeles Times, 3 March 2000.
- [11] New York Times, 2 March 2000.
- [12] Schickler, M.A., Mazer, M.S., and Brooks, C., Pan-Browser Support for Annotations and Other Meta-Information on the World Wide Web. Fifth International World Wide Web Conference, May 6-10, 1996, Paris, France. Internet [September 4, 2002]. WWW: www.5conf.inria.fr/fich_html/papers/P15/Overview.html