Retrofitting Existing Applications and Authoring New Applications for the World Wide Web

Peter D. Karp
Suzanne M. Paley
Artificial Intelligence Center
SRI International, EJ229
333 Ravenswood Ave.
Menlo Park, CA 94025
fax: 415-859-3735
pkarp@ai.sri.com

The CWEST environment allows existing graphical user interface (GUI) applications to be adapted to run through the World Wide Web (WWW) with minor modifications. The same modified application program can run as a stand-alone program on the local window system, or via the WWW. Our group employed CWEST to retrofit the EcoCyc program, which provides visualizations of metabolic pathways and genomic maps, to run through the WWW. CWEST can be used only with GUIs authored using CLIM (Common Lisp Interface Manager). CWEST and CLIM also provide a powerful environment for authoring new WWW applications.

1 Introduction

The advent of the WWW presents both opportunities and challenges. Information providers now have the means of making their information available to a far larger audience than they had previously thought possible. The challenges are two: Can we retrofit existing applications to operate over the WWW without the need for a major rewrite, and to maintain compatibility with the original window system? And can we develop powerful environments for authoring new WWW applications?

We faced this challenge in the context of the EcoCyc GUI [1], which allows users to query and visualize metabolic and genomic data. Potential users must invest significant time and energy to download and install EcoCyc on their computers. In addition, EcoCyc is written using CLIM, which currently runs only on Unix workstations under X-Windows (a PC port of CLIM is underway by Franz Inc.). The prospect of making EcoCyc easily available over the WWW to any biologist with an internet connection is therefore appealing.

CWEST (CLIM-WEb Server Tool) [2] forms a wrapper around the existing EcoCyc application without interfering with its normal operation. The resulting program is a dedicated EcoCyc Web server, that receives requests (URLs) from WWW clients, and returns HTML text and GIF graphics that closely mimic native EcoCyc graphics. This approach was quick to implement, required few changes to the EcoCyc source code, and proved highly effective. The results can be seen at `http://www.ai.sri.com/ecocyc/browser.html`.

2 CLIM Concepts

CLIM is a sophisticated toolkit for building GUIs. Two CLIM concepts are important to the discussion of CWEST. A *presentation* is an abstract graphical object — an association between a piece of screen text and/or graphics, and an application object. For example, the schematic drawing of a chemical compound may be associated with EcoCyc's internal representation for that compound. An *output record* is a CLIM internal data structure that organizes all display output into a hierarchy of output records. The leaf records are the individual output elements, such as a line of text or a graphics operation. Because presentations are a specialized type of output record, all information about presentations resides within the output record hierarchy. CWEST analyzes the output-record hierarchy generated by the EcoCyc GUI to find presentations (which are converted to image maps), and to differentiate textual from graphical regions of a display.

3 EcoCyc WWW Server Operation

The EcoCyc is creating a knowledge base of *E. coli* biochemical compounds, reactions, enzymes, genes, and metabolic pathways. The EcoCyc GUI generates displays of information about each kind of object in the knowledge base. These displays contain part textual information — facts and descriptive comments about an object — and part graphical information, such as chemical structure drawings or metabolic pathway diagrams. The EcoCyc GUI uses a hypertext model to allow users to navigate through the space of metabolic and genomic information. When the user clicks on presentations in an EcoCyc display (such as an enzyme name in a metabolic-pathway drawing), the selected object (the enzyme) is displayed.

Figure 1 outlines the processing of a Web request by the EcoCyc server. Once received, (a) the request is converted to a command to EcoCyc to generate a new drawing (b). The output record tree for that drawing is captured by CWEST (c) and parsed into text regions (d) and graphics regions (e). Text output records are converted directly to HTML (f). A graphics region is converted through several steps to a GIF image (g). Presentations (h) may be embedded within both graphics and text output records, and are parsed to produce hyperlink Universal Resource Locators (URLs) (i). Text hyperlinks can be inserted directly into the HTML output. For graphics hyperlinks, an *image map* (j)

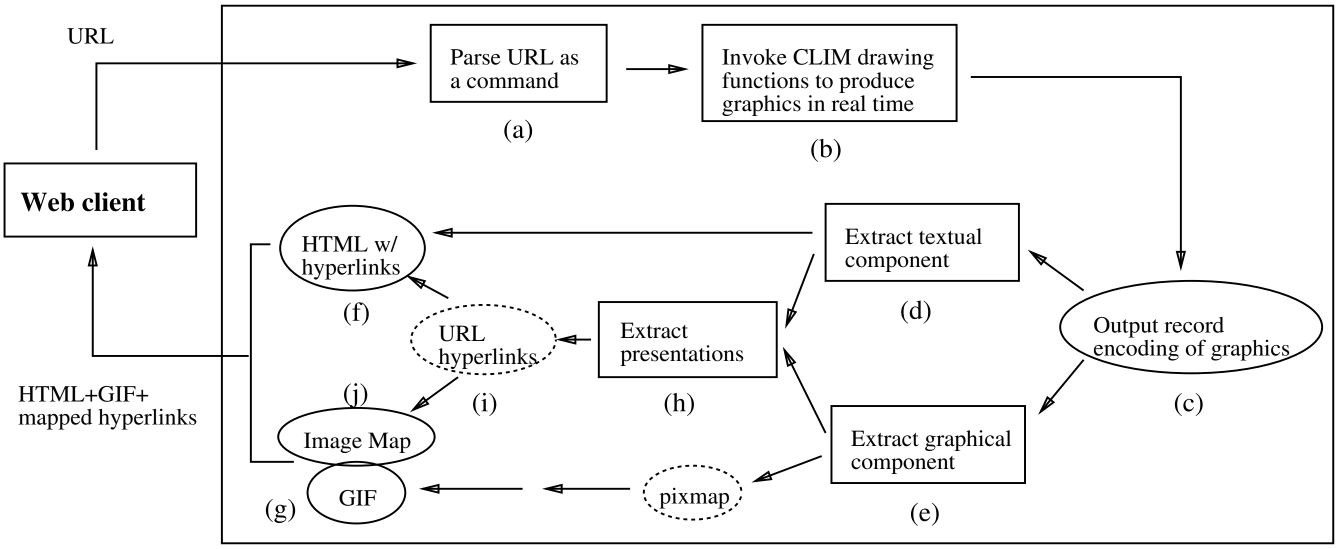

Figure 1: Dataflow diagram illustrating CWEST operation

must be generated that maps each hyperlink to its position in the GIF image. Finally, the HTML output with the mapped inline GIF image is returned to the Web client.

This work was supported by grant 1-R01-RR07861-01 from the National Center for Research Resources.

References

1. P. Karp, M. Riley, S. Paley, and A. Pellegrini-Toole. EcoCyc: Electronic encyclopedia of *E. coli* genes and metabolism. *Nuc. Acids Res.* submitted.
2. S.M. Paley and P.D. Karp. Adapting CLIM applications for use on the world wide web. In *Proceedings of the Association of Lisp Users Meeting and Workshop*, pages 1–9, 1995. See WWW URL `ftp://ftp.ai.sri.com/pub/papers/paley-luv95.ps.Z`.