Accessing Library Catalogs Using Go4zgate

Neophytos Iacovou

While Gopher has had a gateway to WAIS (an extension of the 1988 revision of Z39.50) for several years, Z39.50 has not been standing still. The new 1992 revision of Z39.50 requires a new gopher gateway and this allows for access to online catalogs. This paper describes Go4zgate - a gopher to 1992 Z39.50 gateway.

April 13, 1994

1.0 Introduction

At the moment a student can access the University of Minnesota's library system by either: walking up to any one of the public access terminals found in every library on campus; or they can open a telnet/TN3270 session to the libraries's mainframes. In either case the result is the same, the student is greeted with a what can only be described as "clunky" user-interface.

The University of Minnesota library system has embraced Z39.50 and is beginning to move away from its current mainframe/terminal architecture to a client/server architecture. As this transition occurs the opportunity to give the students a better user-interface exists.

Z39.50 is a NISO standard; its official name is "Information Retrieval Application Service Definition and Protocol Specification for Open Systems Interconnection". Z39.50 is currently at version 2, the planned release date of version 3 is currently set for the fall of 1994.

A Gopher to Z39.50 version 2 gateway (Go4zgate) was developed in order to allow students to access the various Z39.50 databases currently available. At the present time these databases are generally on-line library catalogs, but as more and more people become familiar with Z39.50 the expected number of available Z39.50 servers will increase.

The rest of this paper describes Go4zgate and how it can be used to access on-line library catalogs.

2.0 Why Use Go4zgate?

The benefits of accessing Z39.50 library catalogs via Go4zgate are:

- Users don't have to learn a new IR system, they are already familiar with Gopher.
- The interface is kept simple and does not assume the user to be a "library catalog power user".
- As the user moves from catalog to catalog (potentially from site to site) she is not forced to learn a new command set, the Go4zgate interface remains the same.

3.0 An Example Session

The following is a walk through of what a typical Z39.50 session might look like:

Selecting the desired library catalog to search the user is greeted with a screen asking her to create a query by providing the appropriate data for each of the 3 search fields. At least one of the search fields must be filled in. To keep the user-interface simple all 3 search fields are tied together by a boolean “AND” operation. Go4zgate can also handle an “OR” operation across all fields as well.

The interface to this Z39.50 database creates queries based on the Title, Author, and Subject fields. Had the user not been searching a library catalog but rather a database on cars, the fields from which the query was to be created could just as easily have been Make, Model, and Year. The MaxRecs field prompts the user for the maximum number of records to return from the search.

Being a William S. Burroughs fan this user does a search on Interzone.

FIGURE 1. Creating a Query

```
                 Internet Gopher Information Client v2.0.12

                            ASK block examples

      1.  CC <??>
                                 Melvyl
        Search University of California's Melvyl Library System

   Title     interzone
   Author
   Subject
   MaxRecs   10

                 (enter 1 or more of the search fields)

   [Help: ^_]  [Cancel: ^G]

      15. T-shirt <??>
      16. TextBook.ask <??>
      17. TextBook_Help
      18. Turnip.ask <??>

Press ? for Help, q to Quit, u to go up a menu                Page: 1/2
```

The query returned 5 records. The records are displayed in their “short” form. Selecting any one of the records will display the record in its “long” form.

FIGURE 2. Results of the Query

```
                  Internet Gopher Information Client v2.0.12

                                  Melvyl

 --> 1.  View All Of The Retreived Records: 5 total hits, 5 returned
     2.  Interzone / William S. Burroughs ; edited by James Grauerholz.
     3.  Interzone / William S. Burroughs ; edited by James Grauerholz.
     4.  Interzone : the 1st anthology : new science fiction and fantasy
     5.  Interzone : the first anthology / edited by John Clute, Colin
     6.  Interzone / William S. Burroughs ; edited by James Grauerholz.

Press ? for Help, q to Quit, u to go up a menu                 Page: 1/1
```

Selecting any one of the returned records will display the record in its “long” format. One thing to notice about this format is that information concerning the book’s availability and location within the library are not shown. The reason for this is that holdings information is currently not a part of the Z39.50 standard. This information will be included in version 3 of the standard.

Another thing to note is the “Other authors” field. Z39.50 queries involving the Author are also checked with the “Other authors” field. This may confuse people are first, records will be returned even though the specified author is not the principle author. However, this feature also accounts for anthologies in which the work of the specified author may appear.

FIGURE 3. A Record’s Long Format

```
Interzone / William S. Burroughs ; edited by James Grauerholz. (0k)   100%

Author:        Burroughs, William S., 1914.
Title:         Interzone / William S. Burroughs ; edited by James
                  Grauerholz.
Publisher:     New York, N.Y., U.S.A. : Penguin Books, 1990, c1989.
Pages:         xxiii, 194 p. ; 20 cm.
Notes:         "First published in the United States of America by
                  Viking Penguin ... 1989"--T.p. verso.
Subjects:      Short stories -- lcsh.
Other authors: Grauerholz, James.

[Help: ?]  [Exit: u]
```

The first record returned is actually a pseudo record and is created by Go4zgate. By selecting this record the user is allowed to view the long format of all the records returned. This comes in handy particularly when the result of a query produces more than a handful of records. Without this feature the user would be forced to manually open each and every record in order to gain the same information. Users will also want to save/download this record.

FIGURE 4. All Records Returned

```
View All Of The Retreived Records: 5 total hits, 5 returned (2k)          30%

Author:       Burroughs, William S., 1914.
Title:        Interzone / William S. Burroughs ; edited by James
                 Grauerholz.
Publisher:    New York, N.Y. : Viking, 1989.
Pages:        xxiii, 194 p. : ill. ; 24 cm.
Notes:        Illustrations on lining papers.
Other authors: Grauerholz, James.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Author:       Burroughs, William S., 1914.
Title:        Interzone / William S. Burroughs ; edited by James
                 Grauerholz.
Publisher:    New York, N.Y. : Viking, 1989.
Pages:        xxiii, 194 p. : ill. ; 24 cm.
Notes:        Illustrations on lining papers.
Other authors: Grauerholz, James.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

[Help: ?]  [Exit: u]  [PageDown: Space]
```

This record also supplies the user with some information regarding her request. In this example case the user is told that her query produced 5 records, of which 5 were returned. This feature is useful because often a query may produce more records than were requested by MaxRecs. In such a case the user can then decide if she wants to rerun the query with a higher value for the maximum number of records to be returned.

4.0 Exactly What Information Is Returned?

Currently the following fields are displayed to the user: Author, Title, Publisher, Pages, Series, Notes, Subjects, Other authors, and Call Numbers.

The Call Numbers field can be a bit misleading. Most libraries stopped populating the Call Number field ever since they started to store information about a book in its Holdings Record. Because of this the information in the Call Number field may not always be valid. In the best case there is no information available in the field, or any information present is accurate. In the worst case the information shown in the field is old and out-dated.

In the future, besides returning ASCII text Z39.50 queries may display images or play sounds. These additions are scheduled to find their way into version 3 of the standard.

5.0 Future plans

At the moment the only boolean operation applied is an “AND” or an “OR”, and this is applied across all of the search fields. Future plans involve allowing for boolean operations to be applied within a search field. As an example, the user who previously searched for Interzone would have to issue another query in order to find information about Queer. It would be nice if she could specify something in the order of: “Interzone” AND “Queer”.

We also plan on following the development of Z39.50 and implement changes in the standard in order to keep the gateway as useful as can be.

6.0 Technical Details

6.1 Internet Gopher and the Gopher Protocol

Internet Gopher is an information system used to publish and organize information on servers distributed across the Internet. Initially developed at the University of Minnesota in early 1991, it has spread to over 4800 sites worldwide as of December 1993.

The Gopher system is a client-server system that can be used to build a Campus Wide Information System (CWIS). Clients, which browse and search information are available for most major platforms (Macintosh, DOS, Windows, Unix, VMS, MVS, VM/CMS, OS/2). Servers, which translate and publish information, are also available for all of the platforms mentioned above.

This client-server architecture uses the Internet Gopher Protocol. The Gopher protocol has been described as “brutally simple.” It is based on a web/tree metaphor of files and directories. Its basic primitives are a list directory transaction, a retrieve file transaction and a search for directory entries transaction.

6.2 The Z39.50 Architecture

The last page of this paper contains a diagram illustrating a typical Z39.50 architecture.

The Z39.50 server may have one or more on-line catalogs available. The name of the catalog to be searched must be provided along with the query sent to the server.

Because Z39.50 is a client/server protocol the user's client must make a connection to the Z39.50 server, make a request, and await for the output. Go4zgate was developed because Gopher clients aren't force to speak Z39.50ish; therefore, they require a translator who understands both Gopher+ and Z39.50 version 2. Go4zgate resides between the Gopher+ client and the Z39.50 server.

The interface between the user and Go4zgate is a Gopher+ ASK form. This makes creating new interfaces fast and simple.

The Gopher client, Go4zgate, and the Z39.50 server are all connected together via the Information Super Highway.

6.3 What Does the Ask Form Look Like?

```
Note:          Search University of California's Melvyl Library System
Note:
Ask: Title
Ask: Author
Ask: Subject
Ask: MaxRecs    10
Note:
Note:                (enter 1 or more of the search fields)
%
```

As you can see a rather simple form is used. Note that because MaxRecs must be supplied a value a default number of 10 is supplied for the user. This of course can be changed. If no value is provided by the user the query will not be attempted and the user will be informed that invalid input has been provided.

This interface can be easily expanded as well. If you recall we had previously mentioned that all three of the search fields are tied together by a boolean "AND" operation, and that an "OR" operation was also valid. Because some users may want to decide between the two choices it is possible to add the following line to the above script in order to provide this service:

Choose: Operation: AND OR

In the ASK form shown above this is not an option because it adds another variable which may confuse users. This is not to say that it is never going to be presented to users as an option to control. Below is an example of what a form with more user-specified options might look like:

```
                 Internet Gopher Information Client v2.0.12

                            ASK block examples
                                  Other

                     Search An On-Line Library System

 Title               interzone
 Author
 Subject
 MaxRecs             10
 Boolean Operator    AND

 Hostname of Server  melvyl.ucop.edu
 Catalog to Search   CATALOG
 Port to Connect     210

                  (enter 1 or more of the search fields)

 [Help: ^_]  [Cancel: ^G]  [Cycle Values: Space]  [List Values: l]
     18. TextBook_Help

Press ? for Help, q to Quit, u to go up a menu                Page: 1/2
```

6.4 What Goes on After the User Has Completed the Form?

```
#!/usr/local/bin/perl

$Title = <>; chop ($Title);
$Author = <>; chop ($Author);
$Subject = <>; chop ($Subject);
$Mrecs = <>; chop ($Mrecs);

$Host = "melvyl.ucop.edu";
$Catalog = "CATALOG";
$Port = "210";
$Operator = "AND";
$Gopher_host = "mudhoney.micro.umn.edu";
$Gopher_port = "70";
$MyPID = $$;
$longdir = 0;
$Binary_Path1 = "/export/mudhoney/gopher-data/gplustest/ask/di/go4zgate";
$Binary_Path2 = "/gplustest/ask/di/go4zgate_shell";
$longdir = 1 if ($ENV{"CONTENT_TYPE"} eq "application/gopher+-menu");

system "$Binary_Path1 -I $longdir -h $Host -c $Catalog -p $Port -o $Operator -m
\"$Mrecs\" -T \"$Title\" -A \"$Author\" -S \"$Subject\" -M $MyPID -B $Binary_Pat
h2 -H $Gopher_host -P $Gopher_port";
%
```

The above code is a perl script which processes the information entered into the form. As you can see Operator is preset to an AND operation. Included in this script are preset values for Host (which is set to melvyl.ucop.edu for this particular form), the Catalog, and the Z39.50 port (Port). On thing to note, most servers will be listening to Z39.50 requests on port 210, but this may not always be so.

The only other variable which is interesting to look at is “longdir”. This variable tells Go4zgate if it should return information using the long dir. format or not. As an example, the cursors version of the Gopher client will request for long dir. listings, TurboGopher (the Macintosh Gopher client) will not.

One final word, as you make changes to the ASK form (i.e., by adding options such as allowing the user to select by which boolean operator the search fields are bound) you will also have to make changes to this script.

6.5 A Few Words on Go4zgate itself

You may have noticed that the above script does not apply any error checks on the input provided by the user. All of that is done within the gateway software itself. Rest assured that invalid options are checked for and an error message is returned to the user. Invalid input includes: a value of 0, or less, or NULL for MaxRecs; and NULL data for each of the search fields.

The smaller the catalog the server is searching through and the easier the query the user requested is, the faster the gateway is. The bottleneck in terms of time is, and this shouldn’t come as a big surprise, the time required for the server to execute the query. In order to make life easier for the user, Go4zgate caches the records returned by the server. This means that once the request has been fulfilled the user will be able to select and view any of the returned records relatively fast. If the user tries to retrieve one of the records returned and, for any reason, the cache has been removed the gateway will go back to the server and request that the original query be fulfilled - the original cache will be recreated.

7.0 The Gateway Software

Parts of the gateway software were written by: Ray Larson at the School of Library and Information Studies, UC Berkeley, who wrote most of the Z39.50 query engine; Clearinghouse for Networked Information Discovery and Retrieval (CNIDR), who took the query engine and made a WWW gateway from it; and the University of Minnesota who took the software from CNIDR and made a Gopher+ gateway, in addition to adding more functions provided for by the query engine into the gateway.

The Go4zgate gateway allows the Gopher Client to:

- Request Z39.50 based queries from a Z39.50 server.
- Retrieve and display the records which were returned by the server.

The gateway is available via anonymous ftp from

boombox.micro.umn.edu

via anonymous ftp from the directory

/pub/gopher/Unix/gopher-gateways/go4zgate

Included in this distribution are some sample ASK blocks that define the form users are presented with. Also included is this document.

Architecture diagram showing a Gopher client speaking to a Z39.50 server (with 1 or more on-line catalogs). The Z39.50 Gateway (go4zgate) translates between the two systems.