# Common Gateway Interface

• CGI

•

Copyright (c) 2013-2015 Young W. Lim.
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".
Please send corrections (or suggestions) to youngwlim@hotmail.com.
This document was produced by using OpenOffice and Octave.

# Common Gateway Interface

**CGI** specifies

how the HTTP server interfaces with an external program (gateway)

#### 4 interface components

(a) command line forms of the command line when the server calls the gateway

(b) environment variables these values are transmitted to the gateway

(c) standard input sent to the gateway

(d) standard output sent to the server

# The Output of CGI

# CGI uses standard output

Content-Type: MIME-type

empty line

#### MIME standard content types

- text
- image
- audio
- video
- multipart
- application
- message

#### **HTTP Transactions**

- 1. Connection
- 2. Query
- 3. Processing
- 4. Response
- 5. Termination

#### A query line has 3 parts:

- a query method
- local path of the requested resource
- an HTTP version number

#### A response line has 3 parts:

- an HTTP version number
- a status code
- a textual description of the status

#### **HTML Forms**

#### **Form Handling**

Client side languages: Javascript Server side languages: PHP, JSP

#### form.html

```
<html>
<body>
<form action="form_handler" method="GET">
    User Name: <input name="user" type="text" />
    <input type="submit" />
</form>
</body>
</html>
```

#### form\_handler.php

```
<html>
<body>
<php
// This will print whatever the user put
$name = $_GET['user'];
echo "Hello, ". $name ."!";
?>
</body>
</html>
```

#### form\_handler.cpp

```
<html>
<body>
Hello, world!

</body>
</html>
```

#### **HTTP Transactions**

#### **A CGI script**

- in a special directory: cgi-bin
- obtains URL-encoded query data
- access command-line arguments
- access environment variables
- standard output goes to the server

- through <u>standard input</u>
   for the POST method
- through the QUERY\_STRING environment var for the GET method

- 1. Connection 1. process command-line arguments
- 2. Query 2. obtain form's input data
- 3. Processing 3. process the input data
- 4. Response 4. send a response through a standard output
- 5. Termination 5. terminate

#### Intermediate Nodes

#### **Server specific variables:**

**SERVER SOFTWARE:** name/version of HTTP server.

**SERVER\_NAME**: host name of the server, may be dot-decimal IP address.

**GATEWAY INTERFACE: CGI/version.** 

#### **Request specific variables:**

**SERVER\_PROTOCOL**: HTTP/version. **SERVER\_PORT**: TCP port (decimal).

**REQUEST\_METHOD**: name of HTTP method (see above).

PATH\_INFO: path suffix, if appended to URL after program name and a slash.

PATH\_TRANSLATED: corresponding full path as supposed by server, if PATH\_INFO is present.

**SCRIPT\_NAME**: relative path to the program, like /cgi-bin/script.cgi.

**QUERY\_STRING**: the part of URL after? character. The query string may be

composed of \*name=value pairs separated with ampersands (such as

var1=val1&var2=val2...) when used to submit form data transferred via GET method as defined by HTML application/x-www-form-urlencoded.

**REMOTE\_HOST**: host name of the client, unset if server did not perform such lookup.

**REMOTE\_ADDR**: IP address of the client (dot-decimal).

**AUTH\_TYPE**: identification type, if applicable. REMOTE\_USER used for certain AUTH\_TYPEs.

**REMOTE\_IDENT**: see ident, only if server performed such lookup.

#### Intermediate Nodes

**CONTENT\_TYPE**: Internet media type of input data if PUT or POST method are used, as provided via HTTP header.

**CONTENT\_LENGTH**: similarly, size of input data (decimal, in octets) if provided via HTTP header.

Variables passed by user agent (

HTTP ACCEPT,

HTTP ACCEPT LANGUAGE,

HTTP\_USER\_AGENT,

**HTTP\_COOKIE** and possibly others)

contain values of corresponding HTTP headers and therefore have the same sense.

# Making HTML pages by C++

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

Html.cpp Html.hpp

welcomeTime.cpp

welcomeTime.o: Html.h

welcome: Html.o welcomeTime.o

\$(CC) Html.o welcomeTime.o -o welcome

runwelcome: welcome ./welcome

```
#include "Html.h"
#include <time.h>
int main()
     static time t tod;
     Html page;
     page.leader((char *) "Time to Welcome");
     page.send((char *) "<center><font size=+2><strong>");
     page.send((char *) "Welcome" );
     page.send((char *) "</strong></font></center>");
     page.send((char *) "");
     page.send((char *) "We appreciate your visit.  Our local time is
");
                                                             Content-Type: text/HTML
     page.send((char *) "<em>");
     time(&tod);
                                                             <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 //EN">
                                                  leader()
     page.send(ctime(&tod));
                                                             <HTML><HEAD><rTITLE>Time to Welcome</TITLE>
     page.send((char *) "</em>");
                                                              </HEAD><BODY>
     page.trailer();
     return 0;
                                                             <center><font size=+2><strong>
                                                              Welcome
                                                             </strong></font></center>
                                                              >
                                                             We appreciate your visit.
                                                  send()
                                                              >
                                                              Our local time is
                                                              >
                                                              <em>
                                                             Tue Feb 17 10:50:18 2015
                                                              </em>
                                                              </B0DY>
                                                             </HTML>
                                                 trailer()
```

### leader() and trailer()

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

```
In main()c
//////
           Html.C
                     ///////
#include <fstream>
                                                               page.leader((char *) "Time to Welcome");
#include "Html.h"
using std::ifstream;
                                                                   Content-Type: text/HTML
void Html::leader(char* title)
{ if ( started ) return;
                                                                   <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 //EN">
                                                                   <HTML><HEAD><rTITLE>Time to Welcome</TITLE>
   cout << "Content-Type: text/HTML" << "\r\n\r\n"</pre>
                                                                   </HEAD><BODY>
        << version
        << "<HTML><HEAD><rTITLE>"
        << title << "</TITLE>\n";
   if ( bodyOpen==NULL ) cout << "</HEAD><BODY>\n";
   else sendFile(bodvOpen);
   started = true;
                                                                   </B0DY>
void Html::trailer()
                                                                   </HTML>
{ if ( bodyClose==NULL ) cout << "</BODY>\n";
   else sendFile(bodvClose);
   cout << "</HTML>\n";
                                                     Class Html Constructor
   started = false;
                                                          Html(char* f1=NULL, char* f2=NULL)
                                                           : bodyOpen(f1), bodyClose(f2),
                                                             version((char *) "<!DOCTYPE HTML PUBLIC "</pre>
                                                                               "\"-//W3C//DTD HTML 3.2 //EN\">\n")
                                                          { started = false: }
```

P.S. Wang, "Standard C++ with objected-oriented programming"

### sendFile() and send()

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

#### Read "file" and write it into standard output

```
void Html::sendFile(char* file)
{    if ( !started ) leader((char *) "No Title");
    ifstream_in(file);
    char c[1];
    while ( in.read(c, 1) ) cout.write(c,1);
}

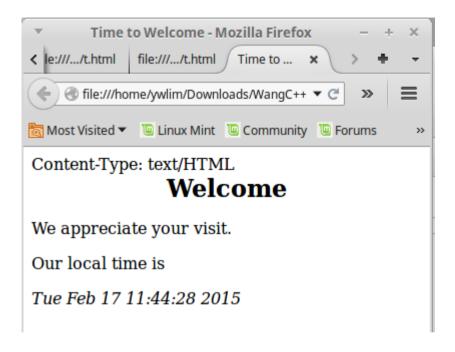
void Html::send(char *s)
{    if ( ! started ) leader((char *) "No Title");
    cout << s;
}</pre>
```

```
////// Html.C
                     ///////
#include <fstream>
#include "Html.h"
using std::ifstream;
void Html::leader(char* title)
{ if ( started ) return;
   cout << "Content-Type: text/HTML" << "\r\n\r\n"</pre>
        << version
        << "<HTML><HEAD><rTITLE>"
        << title << "</TITLE>\n";
   if ( bodyOpen==NULL ) cout << "</HEAD><BODY>\n";
   else sendFile(bodyOpen);
   started = true;
void Html::trailer()
{ if ( bodyClose==NULL ) cout << "</BODY>\n";
   else sendFile(bodyClose);
   cout << "</HTML>\n";
   started = false;
```

#### Class Html header file

```
#ifndef
          Html SEEN
#define
        Html SEEN
//////
           Html.h
                     //////
#include <iostream>
using std::cout;
class Html
{ public:
     Html(char* f1=NULL, char* f2=NULL)
      : bodyOpen(f1), bodyClose(f2),
        version((char *) "<!DOCTYPE HTML PUBLIC "</pre>
                          "\"-//W3C//DTD HTML 3.2 //EN\">\n")
     { started = false; }
     void leader
                    (char* title);
     void trailer
     void send
                    (char *s);
     void sendFile (char* file);
     char* version;
  private:
     bool started;
     char* body0pen;
     char* bodyClose;
};
#endif
```

### HTML File Output



```
ywlim@ywlim-DeskTop-System ~/Downloads/WangC++/ex12 $ ./welcome
Content-Type: text/HTML

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 //EN">
<HTML><HEAD><TITLE>Time to Welcome</TITLE>
</HEAD><BODY>
<center><font size=+2><strong>Welcome</strong></font></center>We appreciate your visit. Our local time is <em>Tue Feb 17 11:44:22 2015

</HTML>
```

# **Decoding URL-encoded Data**

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

Tokenizer.cpp Tokenizer.hpp

testTokenizer.cpp

### testToken1() & testToken()

```
////// testTokenizer.C //////
#include <iostream>
#include "Tokenizer.h"
using namespace std;
                                                                     5=name
                                                                     4=John
void testToken1()
                                                                     3=Doe
            t(" name John Doe age 19 ");
     strina
                                                                     2=age
     Tokenizer tk(t, " "); // " " //
                                                                     1=19
     while ( tk.moreToken() )
     { cout << tk.tokenCount() << "=";</pre>
         cout << tk.nextToken() << endl;</pre>
void testToken()
                                                                     2=name
                t("name=John+Doe"):
     strina
                                                                     1=John+Doe
     Tokenizer tk(t, "="); // "=" //
     while ( tk.moreToken() )
     { cout << tk.tokenCount() << "=";</pre>
         cout << tk.nextToken() << endl;</pre>
```

### testDelTokenizer() & main()

```
16=Test
15=:
14=of
13=:
12=;
11=delimiters
10=:
9=;
8=:
7=also
6=:
5=as
4=:
3=tokens.
2=;
1=:
```

#### Class Tokenizer Header

```
Tokenizer SEEN
#ifndef
#define
           Tokenizer SEEN
//////
          Tokenizer.h
                         //////
#include <string>
using std::string;
class Tokenizer
{ public:
    Tokenizer(const string& str,
               const string& delim = WHITE,
                              want delim = false)
               bool
     : target(str), position(0), delimiter(delim),
       delimToken(want delim) { }
     bool
            moreToken() const;
     string nextToken();
     string setDelimiter(const string& delim) { delimiter = delim; }
            tokenCount() const;
     int
   private:
                            next();
     string
                            nextAll():
     string
     static const string
                            WHITE;
     int
                            position;
     const string&
                            target;
                            delimiter;
     string
     bool
                            delimToken;
};
#endif
```

### moreToken(), nextToken(), next()

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

```
//////
          Tokenizer.C
                        //////
#include "Tokenizer.h"
const string Tokenizer::WHITE=string("\r \n\t");
bool Tokenizer::moreToken() const
{ if ( position == target.length() ) return false;
   if ( delimToken ) return true;
   int i0 = target.find first not of(delimiter, position);
   return(i0 != string::npos);
string Tokenizer::nextToken()
   if ( delimToken ) return nextAll();
                    return next();
    else
string Tokenizer::next()
{ int i0=position;
   int il=position;
                                  // start and end indices
   if ( i0 >= target.length() ) return string();
   i0 = target.find first (not) of(delimiter, i1);
   if ( i0 == string::npos )
                               return string();
   i1 = target.find first of(delimiter, i0);
   position = i1:
   return string(target, i0, i1-i0); // token found
```

string::npos → -1, the end of a string

find\_first\_not\_of(str, pos) → search for
the first character that does not
match any of the characters
specified in str beginning at the
position pos

#### nextAll(), tokenCount()

```
string Tokenizer::nextAll()
{ int i0=position:
   int i1=position;
                                     // start and end indices
   if ( i0 >= target.length() ) return string();
   if ( delimiter.find(target[i0]) != string::npos )
         i1 = i0+1: }
    else
        i1 = target.find first of(delimiter, i0);
         if ( i1 == string::npos ) i1 = target.length();
    position = i1;
    return string(target, i0, i1-i0); // token found
int Tokenizer::tokenCount() const
   if ( position == target.length() ) return 0;
   int count = 0;
   int i0=position, i1=position;
   while (1)
    { if ( i1 == string::npos ) return count;
       i0 = target.find first (not) of(delimiter, i1);
       if ( i0 == string::npos )
           return (! delimToken ? count :
              count + target.length() - i1);
        count++:
       if ( delimToken ) count += i0-i1;
       i1 = target.find first of(delimiter, i0);
```

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

```
FormData.cpp
FormData.hpp
Tokenizer.cpp
Tokenizer.hpp
```

testFormData.cpp

```
FORMOBJ = FormData.o Tokenizer.o testFormData.o
```

```
testForm: $(FORMOBJ)
$(CC) $(FORMOBJ) -o testForm
```

```
runForm: testForm export REQUEST_METHOD=POST; export CONTENT_LENGTH=94; \
./testForm < line
```

```
///////
           testFormData.C
                              ///////
#include <iostream>
#include "FormData.h"
using std::cout; using std::endl;
using std::cin; using std::getline;
int main()
    FormData input;
    if (! input)
    { input.displayError();
       return 1;
    }
                                                                            METHOD = POST
    cout << "METHOD = " << input.getMethod() << endl;</pre>
    cout << "CONTENT = " << input.getContent() << endl;</pre>
    cout << "key count = " << input.keyCount() << endl;</pre>
                                                                            40icm.mcs.kent.edu
                                       cout << input["name"] << endl:</pre>
    if ( input.hasValue("name") )
    if ( input.hasValue("address") ) cout << input["address"] << endl;</pre>
                                                                            kev count = 3
    if ( input.hasValue("email") )
                                       cout << input["email"] << endl;</pre>
    if ( input.hasValue("phone") )
                                       cout << input["phone"] << endl:</pre>
                                                                            Paul S. Wang
                                                                            Math/CS, Kent State U., OH
    return 0;
                                                                            pwang@icm.mcs.kent.edu
// setenv REQUEST METHOD POST
// setenv CONTENT LENGTH 94
// a.out < line
```

```
CONTENT = name=Paul+S%2E+Wang&address=Math%2fCS%
2C+Kent+State+U%2E%2C+OH&email=pwang%
```

```
class FormData
{ public:
     FormData();
    const string& operator [](string key);
    const string& kev(int i) const
    { return (i < name.size()) ? name[i] : empty; }
    int keyCount() const { return name.size(); }
    void enter(string key, string value)
    { data[key]=value; }
    const string& getMethod() const
    { if (error==NO E) return method; }
    const string& getContent() const
    { if (error==NO E) return content; }
    bool hasValue(string key) const
    { return (data.find(key) != data.end()); }
    bool operator !()
    { return error != NO E; }
    void displayError(ostream& out=cerr) const;
    enum ERROR {NO E, METHOD E, LENGTH E, CONTENT E};
```

```
#ifndef __FormData_H_SEEN__
#define __FormData_H_SEEN__
////// FormData.h //////
#include <iostream>
#include <string>
#include "../ex06/token/Tokenizer.h"
#include <vector>
#include <map>
using namespace std;
```

```
private:
                             inputContent();
     ERROR
     void
                             unpack();
     static void
                             urlDecode(string& v); // v decoded
     map<string, string>
                             data;
     vector<string>
                             name;
     string
                             content;
     string
                             method;
     string
                             empty;
     ERROR
                             error;
};
#endif
```

```
/////P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub_bk05.html
          FormData.C
//////
#include <stdlib.h>
#include "FormData.h"
FormData::FormData()
{ content =empty = method = "";
   error = inputContent():
    if ( error == NO E ) unpack();
FormData::ERROR FormData::inputContent()
   char *s = getenv("REQUEST METHOD");
    if (s == NULL) return METHOD E; // no method
    method = s;
   if (method == "POST")
    { s =getenv("CONTENT LENGTH");
      if (s == NULL || *s == '\0') return LENGTH E; // no length
      char* t = s ;
      int len=strtol(s, &t, 10);
      if ( *t != '\0' ) return LENGTH E; // invalid length
       char c;
      while ( len-- > 0 && cin.get(c) ) content += c;
      if (len > 0) return CONTENT E; // missing content
   else if (method == "GET")
    { s = getenv("OUERY STRING");
      if (s == NULL || *s == '\0') return CONTENT E;
       content = s;
    else return METHOD E;
                                       // unknown method
    return NO E;
```

```
int main()
{ FormData input;
```

#### file "line"

```
name=Paul+S%2E+Wang&
address=Math%2fCS%2C+Kent+State+U%2E%2C+OH&
email=pwang%40icm.mcs.kent.edu
```

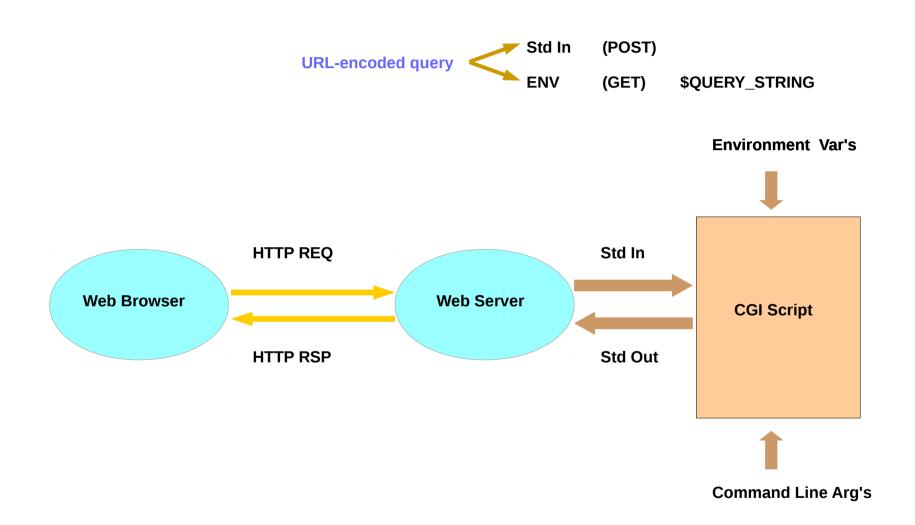
#### in makefile

```
runForm: testForm
   export REQUEST METHOD=POST; \
   export CONTENT LENGTH=94; \
    ./testForm < line
```

```
FormData::FormData()
{    content =empty = method = "";
    error = inputContent();
    if ( error == NO_E ) unpack();
}
```

```
void FormData::urlDecode(string& s)
{ int x1, x2, len = s.size(), pos=0;
    char c, space = ' ';
    for (int i=0; i < len; i++)</pre>
    \{c = s[i]:
        if ( c == '+' ) s[pos++] = space:
        else if ( c == '%' )
        { char x[3] = \{s[++i], s[++i], '\0'\};
           char* t = x:
           int hex=strtol(x, &t, 16);
           if (*t != '\0') // hex seg invalid
           { s[pos++]=c;
               s[pos++]=x[0];
               s[pos++]=x[1];
           else
               s[pos++] = static cast<char>(hex);
        else if ( pos == i ) pos++;
        else s[pos++]=c;
    s.erase(pos,len-pos); // shorten string
```

# HTTP Request and Response



# HTTP Request and Response

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

Html.cpp Html.hpp

welcomeTime.cpp

SUGGESTOBJ = Html.o FormData.o Tokenizer.o suggest.o

suggest: \$(SUGGESTOBJ)

\$(CC) \$(SUGGESTOBJ) -o suggest

runsuggest: suggest export REQUEST\_METHOD=POST; export CONTENT\_LENGTH=253; \ ./suggest < feedback

### **Output Results**

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

```
Content-Type: text/HTML

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 //EN">
<HTML><HEAD><TITLE>Thank You</TITLE>
</HEAD><BODY>
<h3>Thank You</h3>Your message has been logged.We confirm receipt of the following:Subject: Please state shipping charge
From: pwang@icm.mcs.kent.edu <Paul S. Wang>
Date: Fri Feb 20 19:28:37 2015

I am very interested in your home cleaning products.
But there is no shipping charge information.
It would be nice to have such info on your Web site.

</pre
```

Content-Type: text/HTML

#### Thank You

Your message has been logged.

We confirm receipt of the following:

Subject: Please state shipping charge

From: pwang@icm.mcs.kent.edu Date: Fri Feb 20 19:28:37 2015

I am very interested in your home cleaning products. But there is no shipping charge information. It would be nice to have such info on your Web site.

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

```
int main()
    const char
                   *LOGDIR = "./feedback log/"; // dir location
    const int
                    SIZE = 256;
    time t
                    tod:
    ostringstream
                    mystr;
    time(&tod);
    mystr << "./feedback log/" << tod << '\0';</pre>
                    fname(mvstr.str());
    string
    char
                    *logfile = (char *) fname.c str();
   Html
                    page;
    FormData
                    form;
    if ( ! form || ! form.hasValue("subject")
                || ! form.hasValue("email")
                || ! form.hasValue("comment")
       errorReply(page);
        return 1;
                    ofs(logfile, ofstream::out);
    ofstream
    if (! ofs)
    { sorryReply(page);
       return 1;
   enterLog(ofs, form, tod);
   okReply(page, logfile);
    return 0;
```

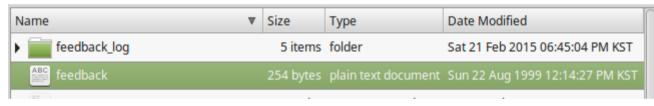
#### Input and Output Files

P.S. Wang, "Standard C++ with objected-oriented programming", http://www.sofpower.com/pub\_bk05.html

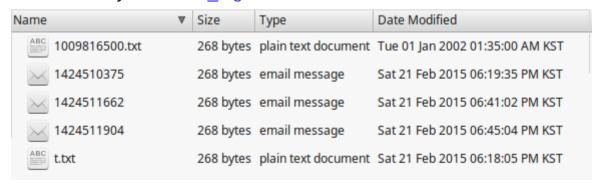
#### file "feedback"

name=Paul+S%2E+Wang&
subject=Please+state+shipping+charge&
email=pwang%40icm.mcs.kent.edu&
comment=I+am+very+interested+in+your+home+cleaning+products%2E++
But+there+is+no+shipping+charge+information%2E+
It+would+be+nice+to+have+such+info+on+your+Web+site%2E

#### Working directory



#### Sub-directory feedback log



### enterLog() & errorReply()

```
//////
                       //////
          suggest.C
#include <iostream>
#include <fstream>
#include <sstream>
#include <string>
#include <time.h>
#include "Html.h"
#include "FormData.h"
using namespace std;
void enterLog(ofstream& out FormData& form, time t tod)
   out << "Subject: " << form["subject"] << endl
        << "From: "
                      << form["email"]
        << " <"
                      << form["name"]
                                         << ">" << endl
                      << ctime(&tod)
        << "Date: "
                                         << endl:
    out << endl
        << form["comment"] << endl;
    out close();
void errorReply(Html& page)
    page.leader((char *) "Form Incomplete");
    page.send((char *) "<h3>Please Complete All Information</h3>"
               "There are missing entries on your form. "
                "Please go back, complete the form, "
                "and submit it again. Thanks.");
    page.trailer();
```

```
enterLog(out, form, tod);
okReply(page, logfile);
```

### sorryReply() & okReply()

```
if ( ! out )
{    sorryReply(page);
    return 1;
}
```

```
enterLog(out, form, tod);
okReply(page, logfile);
```

http://stackoverflow.com/questions/238547/how-do-you-programmatically-download-a-webpage-in-java

```
public static void main(String[] args) {
    URL url:
    InputStream is = null;
   BufferedReader br:
    String line;
    try {
        url = new URL("http://stackoverflow.com/");
        is = url.openStream(); // throws an IOException
        br = new BufferedReader(new InputStreamReader(is));
        while ((line = br.readLine()) != null) {
            System.out.println(line);
    } catch (MalformedURLException mue) {
         mue.printStackTrace();
    } catch (IOException ioe) {
         ioe.printStackTrace();
   } finally {
        try {
            if (is != null) is.close();
        } catch (IOException ioe) {
            // nothing to see here
```

#### Reference

#### References

- [1] http://en.wikipedia.org/
- [2] P.S. Wang, "Standard C++ with objected-oriented programming"