

Chapter 2

RQs + PEs: Accessing Web Resources using URL Connections

Advanced Topics in Java

Khalid Azim Mughal
khalid@ii.uib.no
http://www.ii.uib.no/~khalid/atij/

Version date: 2003-10-01

REVIEW QUESTIONS - URL Connections

1. Which URL components are mandatory in an URL instance? Name the methods to access these components in an URL instance.
2. Explain the methods `openConnection()` in the URL class and the method `connect()` in the `URLConnection` class.
3. Outline the procedure for retrieving a resource via an `URLConnection`.
4. Explain why the following sequence of method calls on a `URLConnection` (denoted by the reference connection) is not correct:

```
connection.getContentType();  
connection.setRequestProperty("User-Agent", "Mozilla/4.0");  
connection.connect();
```
5. How do you obtain an `HttpURLConnection`?

PROGRAMMING EXERCISES - URL Connections

1. Implement a GUI-based program that allows the user to enter a URL. The program should retrieve the source represented by the URL. Only if the content type of the resource is MIME type `text/html`, should the program retrieve all HREF elements and present them to the user in a text area. Some examples of HREF element:

```
<a href="Example1.java">Example 1</a>  
<a href=/advanced_search?q=java&hl=no&lr=&ie=UTF-8&oe=UTF-8>Advanced search</a>  
<a href=http://www.ii.uib.no>My web site</a>  
<a href=http://mannaask.ii.uib.no:8080/index.jsp>Cool Java course</a>
```

Note that there can be several HREF elements in a line and a HREF element can span over multiple lines. As a first approximation you can ignore the case where a HREF element spans over multiple lines.

See the source file `FindingHrefs.java`.