

CS320- Worksheet (2)

REGEX API -JAVA

Write a Java application to read Final Exam Schedule from BC website. The application will read all 4 quarters exams' schedules that are available

URL: <https://www.bellevuecollege.edu/courses/exams/>

Here is a sample run for the expected out of your application:

```
Class Time: 6:30 a.m.
Exam Day: <a href="#spring">See note below</a>
Exam Time: null
+++++
Class Time: 7:30 a.m.
Exam Day: Wednesday, 6/17
Exam Time: 7:30-9:20 a.m.
+++++
Class Time: 8:30 a.m.
Exam Day: Friday, 6/19
Exam Time: 7:30-9:20 a.m.
+++++
Class Time: 9:30 a.m.
Exam Day: Wednesday, 6/17
Exam Time: 9:30-11:20 a.m.
+++++
Class Time: 10:30 a.m.
Exam Day: Friday, 6/19
Exam Time: 9:30-11:20 a.m.
+++++
.....
.....
+++++
Class Time: 9:30 or 10:30 a.m.
Exam Day: Wednesday, 3/24
Exam Time: 9:30-11:20 a.m.
+++++
Class Time: 11:30 a.m. or 12:30 p.m.
Exam Day: Wednesday, 3/24
Exam Time: 11:30 a.m.-1:20 p.m.
+++++
Class Time: 1:30 or 2:30 p.m.
Exam Day: Wednesday, 3/24
Exam Time: 1:30-3:20 p.m.
+++++
Class Time: 3:00, 3:30 or 4:30 p.m.
Exam Day: Wednesday, 3/24
Exam Time: 3:30-5:20 p.m.
+++++
```

Useful information

Reading Directly from a URL

URL is the acronym for Uniform Resource Locator. It is a reference (an address) to a resource on the Internet. You provide URLs to your favorite Web browser so that it can locate files on the Internet in the same way that you provide addresses on letters so that the post office can locate your correspondents.

Sample Java Code:

```
import java.net.*;
import java.io.*;

public class URLReader {
    public static void main(String[] args) throws Exception {

        URL bc = new URL("https://www.bellevuecollege.edu");
        URLConnection yc = bc.openConnection();
        BufferedReader in = new BufferedReader(
            new InputStreamReader(yc.getInputStream ()));

        String inputLine;
        while ((inputLine = in.readLine()) != null)
            System.out.println(inputLine);
        in.close();
    }
}
```

Reference: <https://docs.oracle.com/javase/tutorial/networking/urls/index.html>

Sometime the above code gives forbidden 403 HTML access exception. This is because of *//It's because the site uses SSL.*

You just need to set user agent header for it to work:

SSL (secure socket layer) is a method of ensuring security of data passed back and forth between a client and server. An SSL endpoint is a regular URL, but with https instead of http. Using SSL is more complicated than regular http because there needs to be handshaking between the client and server.

To do that use

```

import java.io.BufferedReader;
import java.io.InputStreamReader;
import java.nio.charset.Charset;
import java.util.regex.Matcher;
import java.util.regex.Pattern;
import java.net.*;

public class URLReader {
    public static void main(String[] args) throws Exception {
        URLConnection bc = new
URL("https://www.bellevuecollege.edu/courses/exams/").openConnection();

        bc.setRequestProperty("user-Agent", "Mozilla/5.0
(Windows NT 6.1; WOW64) AppleWebKit/537.11 (KHTML, like Gecko)
Chrome/23.0.1271.95 Safari/537.11");

        BufferedReader in = new BufferedReader(new
InputStreamReader(bc.getInputStream(), Charset.forName("UTF-
8")));

        String inputLine = "";

        String text = "";
        while ((inputLine = in.readLine()) != null) {
            text += inputLine + "\n";

        }
        in.close();
    }
}

```

// Please note this code is available for you on Canvas to start the activity