

Timed Lab 3 – Class Finder

This Timed Lab is worth 22 **Exam** points. You **MUST** complete the timed lab in your recitation section, physically present with your grading TA for the entire lab, to receive any credit. You may not modify your t-square submission after you leave the timed lab **OR YOU WILL RECEIVE A ZERO**.

<p>For this Timed Lab, you <i>may</i> use</p> <ul style="list-style-type: none">• Course notes• Homeworks• Recitation assignments• Other course material• Any material you may find on the Internet that don't involve communicating "live" with other people.	<p>However, you <i>may not</i></p> <ul style="list-style-type: none">• Communicate with other people/students in real-time via any means. This means no Facebook, email, new Piazza posts, IM, IRC, cell phones, Google Talk, smoke signals, etc.• Share code with other students.• Look at other students work.
--	--

The TAs will be available to answer clarifying questions about the problem, but they are not permitted to give assistance with debugging code, program logic, etc. You will have an entire recitation period to work on this assignment; this time begins *exactly* when your recitation begins and ends *exactly* when your recitation ends: No extra time will be given if you arrive late, except in highly extenuating circumstances that must be approved by Dr. Summet.

T-Square will not permit any late submissions; ensure that you submit your code to T-Square several times to prevent earning a zero due to you being unable to submit. Your TAs will give a verbal warning 10 and 5 minutes before the end of the recitation period; you should submit at these times.

In your collaboration statement, if you use code from somewhere that is *not* a class resource (i.e. not listed on the course calendar), please list where this code came from. Ensure that you fill out the header at the top of the file.

Problem Description:

In this timed lab, you will be scraping a webpage for information about a particular class that will be meeting next semester. You will build a GUI that will allow the user to scrape the page and display all of the course listings in the GUI.

Website Format & Location:

You will scrape the website located at http://www.cc.gatech.edu/classes/AY2012/cs2316_fall/tl3.html The website format is extremely simple HTML; the only components you need to worry about are the title of the webpage, contained within the <title></title> HTML tags, which specifies the name of the course, and a list of course times (Start – End), specified in an HTML ordered list. The list will be formatted as follows:

```
<ol>
<li>HH:MM – HH:MM SECTION_ID
<li>HH:MM – HH:MM SECTION_ID
...
```

HH:MM – HH:MM SECTION_ID

You will be required to find the start time, end time, and Section ID for each section from the webpage. If you choose to use regular expressions to find all of the sections, here is some more information which may be of use:

HH can range from 00 to 23. MM can range from 00 to 59. There is a space, then a dash, then a space between the two different HH:MM's. There are two spaces between the ending time of the class and the SECTION_ID. The SECTION_ID is at least one, but possibly more letters. The use of regular expressions is not required, but could make this problem much simpler.

A second testing website with the same format but slightly different data is located here:

http://www.cc.gatech.edu/classes/AY2012/cs2316_fall/tl3a.html

you may use either website URL for your program, but your program must work with any webpage that follows the specified format, not just the two example webpages. Different classes may have different numbers of sections.

The GUI:

This GUI will be of your own design, but must incorporate a few elements at a minimum.

- There must be a button that when clicked will scrape the webpage and display the relevant information in the GUI. *After the button is clicked, you must set the button state to disabled so the user cannot click it again.*
- You must display in a widget the name of the class (from the webpage title)
- Each section from the class must be displayed in its own widget. You can split up the start time, end time, and section ID into three widgets per section if you wish to lay your GUI out this way, but you must have at minimum one widget in the GUI per section in the webpage.
- There is no formatting requirement on how the information should be displayed in the GUI, except for the above requirements.

Make sure to add the required code to make your GUI appear when your file is run!

Grading:

+2 – **Displays GUI when file is run**

+5 – **GUI setup**

+1: Saves rootwin reference so widgets can be added after button is pressed.

+2: Contains at least the button initially.

+2: Clicking the button causes the webpage to be scraped. (Using either URL)

+10 – **Web Scraping**

+3: Correctly downloads webpage

+2: Correctly finds title of webpage

+5: Correctly finds each section of the course

+5 – **GUI Updating**

+3: Correctly places each section's information into the GUI

+1: At least one GUI widget per section.

+1: Button disabled after webpage is scraped.