

Timed Lab 3 – E-mail Collector

This is a Timed Lab; this Timed Lab is worth 22 **Exam** points.

<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, Piazza, 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.

Note that you must check out with your TA before you leave the recitation room. If you do not check out with your TA or you modify your submission after you leave the recitation room, **you will receive a grade of zero on the timed lab**. No submissions will be accepted after T-Square closes the assignment (i.e. it will not let you submit).

Problem Description:

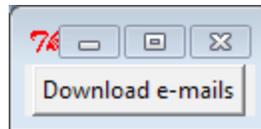
In this timed lab, you will be scraping a TA Helpdesk file for the TAs e-mail addresses so you can e-mail them with questions about your homework!

Website Format & Location:

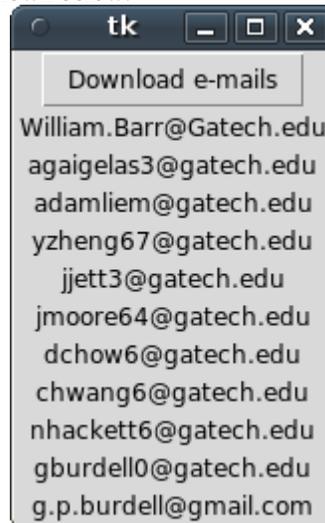
You will scrape the website located at:

http://www.cc.gatech.edu/classes/AY2012/cs2316_spring/codesamples/helpdesk.html

This helpdesk file is based upon the helpdesk file for CS 2316 this semester, with a few minor alterations and additions. Your objective is to retrieve all e-mail addresses and place them in a simple GUI, which may (but does not need to) look something like below (it must have all the same elements as the one shown):



Clicking on the Download e-mails button will download the webpage, scrape the e-mail addresses, and place them in the GUI in some fashion, like the one shown below:



Note that all e-mail addresses have had their [dot] and [at] placeholders converted to the proper symbols, and the extraneous spaces in the middle of the e-mail addresses have been removed, producing e-mail addresses that are suitable to send e-mails to.

The E-mail Formats:

E-mail addresses generally follow the format RECIPIENT@DOMAIN.TLD (TLD stands for Top Level Domain). Your program should be able to handle the following for each portion of the address:

- The . is interchangeable for “ [dot] ” and the @ is interchangeable for “ [at] ” (note the spaces in both) at any point in the e-mail.
- RECIPIENT must start with a letter (upper or lower case) or digit and then may be followed by any number of letters or digits or dots.
- DOMAIN will be any number of letters (upper or lower case) or digits.
- TLD will be any number of letters (upper or lower case) or digits.

Grading:

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

+5 – **GUI contains all required elements and correctly adds widgets when button is clicked**

+2: Contains “Download e-mails” button.

+1: Adds widgets to GUI when button is pressed.

+2: Widgets with any type/number of email addresses are added when button is pressed.

+10 – **Web Scraping / Data Finding**

+3: Correctly downloads webpage.

+7: Correctly finds each e-mail address as specified in email formats above.

+5 – **E-mail Conversion**

+3: Correctly replaces [dot] and [at] with appropriate characters

+2: Extraneous spaces removed from e-mail addresses before being added to the GUI.