Data Structures for Data Science

WHAT’S IN THIS SYLLABUS?
- Course Description
- Learning outcomes
- Policies and Expectations
- Assignment Descriptions

COURSE INFORMATION
- 4 Credit Hours (Undergraduate)
- Monday 3:35-4:25pm (Hardman Hall 101)
- Tuesday & Thursday 3:30- 4:45pm (Dawson Hall 306)
- CSCI2725; Spring 2020

INSTRUCTOR INFORMATION
- Soheyla Amirian (Ph.D. Candidate)
- Office: 223 Boyd
- Office hour: Tuesday 2:30-3:30pm or by appointment
- Email: amirian@uga.edu

COURSE OVERVIEW
This course surveys data structures and explores their different implementations with a focus on their manipulating algorithms. The objective is to effectively use data structures and their manipulating algorithms to design and implement efficient solutions and to develop problem-solving skills required in the data science paradigm.

The course introduces approaches to algorithm design, including divide and conquer, greedy algorithms, and dynamic programming and will draw on applications from data science.

LEARNING OUTCOMES
1. **Use** asymptotic analysis notations to give upper bounds on time and space complexity of algorithms.
2. **Design, analyze, and implement** recursive functions.
3. **Describe, design, and implement** generic, reusable abstract data types (ADTs), including sorted list, unsorted list, tensors, stack, queue, priority queue, tree, graph, and dictionary.
4. **Analyze** and assess the impact of data structures and algorithm design on the performance of algorithms.
5. **Understand** algorithm design methods such as the greedy method, divide and conquer, and dynamic programming.
6. **Choose** the appropriate data structures and algorithm to solve a real-world problem and to defend the selection.
7. **Design, apply, and compare** the complexity of principal algorithms for sorting, searching, and hashing.
8. **Understand** and apply graph algorithms, such as graph traversal algorithms, minimum spanning tree, topological sort, and single-source shortest path.
9. **Write** programs to satisfy the requirements of a real-world problem, integrating course concepts (data structures and algorithms efficiency) and object-oriented design principles.
TOPICAL OUTLINE
1. Algorithm Analysis
2. Fundamental Data Structures
3. Recursion
4. Binary Search Trees (BST), and Balanced Trees.
5. Heaps and Priority Queues
6. Understand, Design, and implement the Priority Queue ADT using various data structures
7. Graphs and Graph Algorithms
8. Algorithm design strategies
9. Hash Tables
10. Implement and assess sorting algorithms
11. Discuss factors other than computational efficiency that influence the choice of algorithms
12. Understand and use advanced data structures for data science, including tensors, disjoint sets, data frames, and decision trees

Recommended Books
Author: Michael T. Goodrich, Roberto Tamassia, Michael H. Goldwasser
Or
Data Structures and Algorithms in Python; by Goodrich, Tamassia, Goldwasser

Prerequisite: CSCI 1302
Reference Materials
Additional texts and notes may be suggested for reading throughout the semester. If your instructor posts something for you to read, you are expected to read it and try to understand it, even if it's not part of an official assignment.

COMMUNICATION
Please check your UGA email address and eLC frequently. All emails to the instructor must come from your UGA email address and must include “CSCI2725" in the subject line. Please allow 1 business day for the instructor to respond to your email.

Your expectations of me as instructor and facilitator:
- Provide clear expectations for assignments and lectures
- Be flexible and understanding

Expectations for yourselves:
- Be on time and prepared
- Think big
- Be productive
- Responsive and provides feedback
- Be open to new perspectives and ways of thinking
- Be patient
- Be active and involved in the class

**POLICIES**

**Honor Code and Homework**

UGA Student Honor Code: "I will be academically honest in all of my academic work and will not tolerate academic dishonesty of others." A Culture of Honesty, the University's policy and procedures for handling cases of suspected dishonesty, can be found at www.uga.edu/ovpi. Every course syllabus should include the instructor's expectations related to academic integrity. You may discuss the problem and solution strategies with your classmates but the work you turn in has to be yours and should reflect your effort.

Teamwork is allowed for projects.

Late paper review will be accepted with 10% reduction points for each day.

**Electronic Devices**

Please silent your devices before class. Laptops and tablets may be used only for class activities.

**Accommodations**

I will work with each of you to maximize your learning opportunity. If you anticipate issues related to the format or requirements of this course, please meet with me. I would like us to discuss ways to ensure your full participation in the course. If you determine that formal, disability-related accommodations are necessary, it is very important that you be registered with the Disability Resource Center. They can be reached by visiting Clark Howell Hall, (by calling 706-542-8719 (voice) or 706-542-8778 (TTY) or by visiting http://drc.uga.edu) and notify me of your eligibility for reasonable accommodations. We can then plan how best to coordinate your accommodations.

**ASSESSMENT**

<table>
<thead>
<tr>
<th>Section</th>
<th>Portion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>10%</td>
</tr>
<tr>
<td>Quiz &amp; Activities*</td>
<td>15%</td>
</tr>
<tr>
<td>Team Projects**</td>
<td>35%</td>
</tr>
<tr>
<td>Midterm and Final exam</td>
<td>40%</td>
</tr>
</tbody>
</table>

**FINAL LETTER GRADES**

<table>
<thead>
<tr>
<th>A ≥ 93</th>
<th>93 &gt; A- ≥90</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 &gt; B+ ≥87</td>
<td>87 &gt; B ≥83</td>
</tr>
<tr>
<td>83 &gt; B- ≥80</td>
<td>80 &gt; C+ ≥77</td>
</tr>
<tr>
<td>77 &gt; C ≥73</td>
<td>73 &gt; C- ≥70</td>
</tr>
<tr>
<td>70 &gt; D ≥60</td>
<td>F &lt; 60</td>
</tr>
</tbody>
</table>

*Activities are the in-class questions and participation that have bounce points.

**I recommend team project of 2 or 3 people.

Midterm Exam on February 28th, Class time
Final Exam on Thursday, April 30th, on 3:30pm to 6:30pm