Syllabus

COSC 373: Distributed Algorithms

Will Rosenbaum
Department of Computer Science
Amherst College
Updated: February 2, 2022

1 Basic Information

Instructor Will Rosenbaum, SCCE C216

Lectures T/Th 10:00–11:20, SCCE A131

Office Hours TBD

Course Website willrosenbaum.com/teaching/2022s-cosc-373/

2 Official Course Description

A distributed system consists of a network of processors that communicate by exchanging messages. No processor has a global view of the network, so neighboring processors must communicate in order for the system to perform a given task. In this course, we will study the theory of distributed systems. We will consider fundamental algorithmic tasks—for example, finding spanning trees, maximal independent sets, and graph coloring—in several models of distributed computing. Our goal is to understand under what conditions these tasks can be performed efficiently, if at all. While this course is primarily theoretical, we will discuss applications of the theory to modern computing paradigms (e.g., MapReduce) as time allows.

3 Course Aims

The aim of this course is to lay the algorithmic and formal foundations for reasoning about distributed systems. Broadly speaking, a distributed system is any system of autonomous entities that work in concert to perform some task. Computer networks are probably the most obvious example of distributed systems, but there are many other examples of distributed systems in nature as well: social networks, ant colonies, and the human brain are just a few.

Given the diversity of distributed systems, there are *many* different computational models that highlight various features of distributed systems. We will consider several computational models in order to understand their relative power and limitations. As we will see, small changes in the model can have a large impact on the capabilities of a system. Thus, a primary goal of the course is for you to develop *formal* reasoning skills to address these subtleties.

Finally, this course will provide a gentle introduction to reading primary literature (i.e., research papers) in computer science.

4 Expected Background

You are expected to be familiar with basic algorithmic techniques and the analysis of algorithms. You should be comfortable with "big O" notation as it applies to bounding the running times of algorithms. Additionally, you should be familiar with graphs and basic graph algorithms such as breadth- and depth-first search, finding shortest paths (e.g., Dijkstra, Bellman-Ford, and Floyd-Warshall), and minimal spanning trees (e.g., Prim and Kruskal). Some familiarity with NP-completeness and reductions as well as basic proof techniques (e.g., induction and proof by contradiction) will be helpful, but not strictly necessary.

5 Resources

There is no single text that covers all of the material we will discuss in this class. The lectures will draw from several textbooks as well as primary literature. Required and relevant readings will be posted to the course webpage. The sources listed below cover most of the material from the class (and much more material that we will not cover). Most of the available texts on distributed algorithms are written primarily for a graduate-level courses, and thus are not especially easy reads. Much of class time will be devoted to discussing the assigned readings.

Main Texts

- *Distributed Algorithms* by Hirvonen and Suomela (available online)
- *Distributed Network Algorithms* by Pandurangan (available online)
- *Principles of Distributed Computing (Lecture Notes)* by Wattenhofer (available online)

Other Recommended Texts

- *Distributed Algorithms* by Lynch (Chapters 1 through 4)
- *Distributed Computing: A Locality-Sensitive Approach* by Peleg

Other resources will be posted to the course website when then are relevant to class discussion.

6 Topics Covered

The course will cover the following topics:

1. stable and maximal matchings in the Port Ordering (PO) model
2. impossibility results in the PO model
3. global constructions in the CONGEST model
    - leader election and breadth-first search
    - minimum spanning trees
    - diameter and all-pairs-shortest-paths
4. 2-party communication complexity
5. CONGEST lower bounds
6. local verification and proof-labeling schemes
7. randomized distributed algorithms and maximal independent sets

Additional topics may include asynchronous computation, symmetry breaking in the LOCAL model, and efficient algorithms via all-to-all communication.

7 Coursework and Evaluation

The primary coursework will consist of:

- written homework assignments ($\sim 30\%$)
- take-home quizzes ($\sim 20\%$)
- a mid-term exam ($\sim 20\%$)
- a final group project ($\sim 30\%$)

Homework and quizzes will be assigned on a bi-weekly basis, due on alternating weeks. Homework assignments and the final project can be done in small groups; quizzes and the mid-term exam are to be completed individually. Final grades will take into account the four categories above, with the indicated (approximate) weights.

All coursework will be posted to the course Moodle site and Gradescope page, and should be submitted through Gradescope. **Late work will not be accepted without prior approval.**

8 Course Policies

Class participation. While attendance is not strictly mandatory, you are expected to actively participate in class on a regular basis. Participation includes both preparing for class by engaging with the assigned readings as well as contributing to classroom discussion.

COVID protocols & absences. You are expected to follow college-wide COVID precautions. In particular, you should wear appropriate masks ((K)N-95 or KF-94) at all times in class, and there is no eating or drinking in the classroom. **If you are under isolation/quarantine or are feeling ill, you should not attend class.** Class/lecture notes will be made available so you will have an idea of what was discussed in class.

Collaboration. Collaboration is encouraged on homework assignments and the final project. For these group assignments, all students in the group are expected to contribute equally.

Take-home quizzes and the midterm exam are to be completed individually. You should not discuss the contents of quizzes/exam with any person before all students have submitted their work.

Academic Integrity. Your expected to uphold the Amherst College Honor Code. In particular, you should not discuss material appearing on a quiz or exam before you or others have submitted the test, or submit work either done by someone else or derivative of someone else's work without crediting the original author/source.

Failure to comply with these terms may result in loss of credit for the assignment or course, as well as possible disciplinary action by the college.