Arizona State University

CSE 598 SLN 89535 — **Markov Chain and Monte Carlo Methods** — Fall 2019

Session C: August 22nd – December 6th, 2019


| | |
|---|---|
| **Instructor:** | Joshua J. Daymude (`jdaymude@asu.edu`, BYENG 498) |
| **Co-Instructor:** | Andréa W. Richa (`aricha@asu.edu`, BYENG 440) |


| | |
|---|---|
| **Lectures:** | MW 4:35pm–5:50pm, BYENG 210 |
| **Office Hours:** | W 10:00am–11:00am, Th 3:00pm–4:30pm |
| **Website:** | `https://joshdaymude.wordpress.com/teaching/asu-f19-mcmc/` |

## Course Description

Markov chains and Monte Carlo methods form the foundation of a cohesive and elegant theory aimed at (*i*) counting the size of very large sets — e.g., the set of all matchings of a large graph — and (*ii*) randomly sampling from such sets. This seminar aims to provide a deep understanding of discrete-time Markov chains and their applications from a theoretical perspective, focusing primarily on problems in graph theory, combinatorics, statistical physics, computational geometry, and optimization. The seminar is as follows.

We begin with one of the few known polynomial-time algorithms for *exact counting*; i.e., exactly counting the size of a large combinatorial object. However, this is immediately followed by a negative result that shows such problems are, in general, **#P**-complete. This limitation motivates the need for methods achieving *approximate counting*, which is tightly linked to *random sampling*. Here, we introduce Markov chains as well-studied tools for achieving these goals. Markov chain fundamentals such as stationary distributions are covered, and several simple example chains are described. This completes the first unit of the seminar.

The second unit is concerned with techniques for analyzing the *mixing time* of a Markov chain, which is a critical part of showing that the resulting Monte Carlo methods are efficient. We will cover the coupling technique, conductance arguments, spectral gap techniques, canonical paths, and coupling from the past. More example Markov chains and analyses will be given along the way to bolster student familiarity with Markov chains.

Finally, the seminar concludes with a series of Markov chain/Monte Carlo method applications, including applications in graph theory, statistical physics, phylogeny, volume computations, and Bayesian inference. These topics may change based on time constraints and student interest.

## Prerequisites

As a topics seminar, Markov Chains and Monte Carlo Methods is intended for a diverse audience from a variety of academic disciplines. While the material has broad appeal even outside the computational sciences, it should be especially interesting to those in theoretical computer science or mathematics working in discrete mathematics, graph theory, probability theory, and randomized algorithms. However, as much as the lectures will aim to be self-contained, this course will be very difficult for those who have not had some introduction to formal logic, mathematical reasoning, and the design and analysis of algorithms. Interested students should consider the following prerequisites as guidelines.

1. **At Minimum:** A desire to learn complex material, and a willingness to ask questions!
2. **Baseline:** Exposure to some form of symbolic logic or proof-based mathematical reasoning (e.g., PHI 333, MAT 243, MAT 300), or a very strong grasp of the theoretical aspects of introductory algorithms courses (e.g., CSE 310, CSE 450/551).
3. **Preferred:** Some grasp of probability theory and its distinction from statistics (i.e., not IEE 380, but STP 326, STP 421, STP 501). Some understanding of graph theory (i.e., MAT 416/513) and its relationship to computer science. Experience with the design and analysis of algorithms (e.g., CSE 450/551) and a loose understanding of complexity classes (CSE 550 and CSE 555).
4. **Well Prepared:** Understanding of advanced graph theory (MAT 516 and MAT 517), combinatorics (MAT 515), randomized and approximation algorithms (CSE 552), and probability theory and stochastic processes (IEE 640, STP 425, STP 540).

## Textbooks

The notes available on the course site will serve as the primary resource, but the following texts are excellent references:

1. *Counting, Sampling, and Integrating: Algorithms and Complexity* by Jerrum [1].
2. *Markov Chains and Mixing Times* by Levin-Peres-Wilmer [2].

## Evaluation

Students will be evaluated according to the following rubric:

| | |
|---|---|
| **Participation** | (15%) |
| **Problem Sets** | (30%) |
| **Final Project Report** | (25%) |
| **Final Project Presentation** | (30%) |

The problem sets and final project report/presentation must be submitted on Canvas by the posted due dates to receive credit. Students are encouraged to discuss any personal circumstances that may hinder timely completion of work with the instructor ahead of time, if possible, and extensions may be granted in rare cases at the instructor's discretion.

There will be no exams, and attendance is not mandatory. However, as this is a topics seminar, it is important to show up and participate in lectures in order to learn the material. If a student misses a lecture, it is their responsibility to attend office hours and review the notes to catch up.

**Participation.** Based on a student's involvement in lectures and office hours. This is not intended to punish students who prefer to be quieter in class, but to encourage all students to seek clarification on things that are unclear or subtle (of which there will likely be many, as this material can be quite complex).

**Problem Sets.** Three to four problem sets will be assigned throughout the seminar based on material covered in class. Students are encouraged to seek help from the instructor during office hours after carefully reviewing the relevant material. Solutions will be discussed in class after the assignments are due.

**Final Project.** Students must complete *one* of the following types of final projects. No type will be graded more favorably than any other, despite possible differences in difficulty between them.

1. **Implementation:** Students will find a suitable research paper utilizing Markov chains and/or Monte Carlo methods and will perform a rigorous validation of the results therein using an original implementation. Preferably, the students will extend the results of the paper to new application domains or datasets, designing and executing their own experiments. Code can be written in a language of the student's choosing, but must be well documented (both with in-code comments and external documentation, if necessary).

2. **Survey Paper:** Students will choose a technique or application area related to Markov chains and Monte Carlo methods and will write a rigorous survey paper on it. The survey must be unified by a specific line of questioning or investigation; it should not simply list a long summary of various papers. The final paper must cite at least 20 relevant references.

3. **Original Research:** Students will pose an original research question relating to or using Markov chains and/or Monte Carlo methods and will write a paper investigating the research question. The format of a standard academic paper for the topic area is expected, including an introduction, related work, etc.

The final project is split into three phases. Students will first propose a project topic in mid-October. There will then be a roughly 6-week period to carry out the project and prepare a report on their findings. Finally, students will give in-class presentations in the last two weeks of the semester in the format of a conference talk (roughly 10–20 minutes per person, depending on the number of students). Students may work alone or in pairs. In the case of a two-person group, each group member will receive the same grade on their project report but may receive individual grades for their presentation.

## Classroom Behavior

The instructor and students should treat one another with kindness and respect. All are responsible for maintaining a learning environment with minimal distractions so all can remain present and engaged. Any disruptive, threatening, or violent behavior will be reported directly to the ASU Police Department and Office of Student Rights and Responsibilities.[1]

## Communication

The instructor will communicate with the class via announcements on Canvas and/or updates to the course website,[2] and will respond to emails and Canvas discussion board posts daily during

[1] SSM 104–02: Handling Disruptive, Threatening, or Violent Individuals on Campus. `https://www.asu.edu/aad/manuals/ssm/ssm104-02.html`

[2] MCMC Course Website. `https://joshdaymude.wordpress.com/teaching/asu-f19-mcmc/`

business hours. Students are encouraged to bring more in-depth questions and concerns to the instructor after lectures or during office hours. Emails to the instructor (`jdaymude@asu.edu`) must have a subject line beginning with `CSE 598:`, and the body of the email should be professional, clear, and concise. Discussion boards will be available on Canvas for questions and concerns that are relevant to the whole class.

## Academic Integrity

Students in this class must adhere to ASUs academic integrity policy.[3] Students are responsible for reviewing this policy and understanding each of the areas in which academic dishonesty can occur. In addition, all engineering students are expected to adhere to both the ASU Academic Integrity Honor Code[4] and the Fulton Schools of Engineering Honor Code.[5] All academic integrity violations will be reported to the Fulton Schools of Engineering Academic Integrity Office (AIO). The AIO maintains record of all violations and has access to academic integrity violations committed in all other ASU college/schools.

Course content — including lectures, notes, and assignments — are copyrighted materials that students may not share outside the class, upload to online websites not approved by the instructor, sell, or distribute. Students must obtain verbal consent from the instructor before recording lectures, and such recordings are for personal use only.[6] The Fulton Schools of Engineering reserve the right to delete materials found online on the grounds of suspected copyright infringement.

## Disability Accommodations

Students seeking accommodations should register with the ASU Disability Resource Center[7] and request accommodations through their forms. Every effort will be made to comply with these accommodations to foster a productive learning environment.

## Discrimination, Harassment, and Retaliation

Arizona State University is committed to providing an environment free of discrimination, harassment, or retaliation for the entire university community, including all students, faculty members, staff employees, and guests. ASU expressly prohibits discrimination, harassment, and retaliation by employees, students, contractors, or agents of the university based on any protected status: race, color, religion, sex, national origin, age, disability, veteran status, sexual orientation, gender identity, and genetic information.

Your safety is paramount. If you or someone you know has been sexually harassed or assaulted, ASU encourages you to consult the ASU Sexual Violence FAQs[8] for information and resources. Please be aware that many university officials (including the instructor) are obligated to report

[3] ASU's Academic Integrity Policy. `https://provost.asu.edu/academic-integrity/policy`

[4] ASU's Academic Integrity Honor Code. `https://provost.asu.edu/academic-integrity/honor-code`

[5] Fulton Schools of Engineering Honor Code. `https://engineering.asu.edu/ira-a-fulton-schools-of-engineering-honor-code/`

[6] ACD 304–06: Commercial Note-Taking Services. `https://www.asu.edu/aad/manuals/acd/acd304-06.html`

[7] ASU Disability Resource Center. `https://eoss.asu.edu/drc`

[8] ASU Sexual Violence FAQs. `https://sexualviolenceprevention.asu.edu/faqs`

any incidents of discrimination, harassment, or assault they become aware of to the Offices of Student and University Rights and Responsibilities. If you would instead like to report an incident confidentially, trained professionals at ASU Health and Counseling Services are available to you.

## References


[1] M. Jerrum. *Counting, Sampling, and Integrating: Algorithms and Complexity*. Lectures in Mathematics. ETH Zürich. Birkhäuser, April 2003.

[2] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2nd edition, October 2017. Available online at `http://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf`.