Information Theory in TCS (COMS-E6998, Fall'17)

Logistics:

Time: T 4:10pm - 6:40pm.
Location: TBD.
Office hours: TBD.

Teaching assistants:

- TBD

Description:

The goal of this course is to develop tools in information theory and communication complexity for understating computation, mostly to prove unconditional lower bounds on various computational models such as streaming algorithms, data structures, distributed computing, linear programs, circuit design and even economic scenarios. On the upper bound side, we will discuss compression schemes for one-way and interactive protocols, their implications to theoretical computer science, and the limits and open directions in generalizing Shannon’s classical information theory to the interactive setup (aka Information Complexity). Time permitting, we will discuss some questions in "algorithmic information theory” such as constructing efficiently decodable capacity-achieving codes.

This is an advanced course geared towards CS and EE graduate students, though it is designed to be self contained. Evaluation is based on home works and a final project (reading, implementation, or research).

The class satisfies the track electives for Theory (Foundations) track.

Prerequisites:

There are no mandatory prerequisites other than familiarity with probability theory and linear algebra. Background in Complexity Theory (e.g., COMS 4236) and information theory is recommended but not a prerequisite. However, mathematical maturity is a must, and lectures are based on theoretical ideas and will be proof-heavy. Students are expected to be able to read and write formal mathematical proofs.

Tentative list of topics:

Communication complexity (deterministic, non-deterministic, randomized), applications to linear programs (Yannakakis’ Factorization theorem and the “Clique vs. Independent Set” problem), streaming lower bounds and data structure lower bounds (e.g., near-neighbor search, range counting, dictionaries, polynomial evaluation), Information Theory 101 (definitions and basic inequalities), Shannon’s noiseless coding theorem (and its interactive
analogue), Shearer’s Lemma and applications to computer science, Information Complexity, interactive compression schemes and the Direct Sum conjecture, algorithmic information theory.

List of similar courses taught at other universities:

- Information Theory in Computer Science (Madhu Sudan, Harvard).
- Information Theory and Applications (Anup Rao, UW).
- Communication Complexity (Prahladh Harsha, TIFR).

Course reference books:

- Communication Complexity (E.Kushilevitz, N.Nisan).
- Elements of Information Theory (T.Cover).

Grading:

Grading will be based on bi-weekly home-work assignments (50%), scribing one lecture (10%), and a final project (40%).

Homeworks: A problem set will be handed once every 2 weeks (5-6 in total). Collaborate on home-works is allowed but writing formal solutions should be done *separately*. In this case, you should write clearly all the persons you have collaborated with. There will also be an automatic extension policy. All submissions are via CourseWorks.

The final projects can be of two types:

- **Reading-based**: choose a course-related research paper (approved by the instructor), summarize it in writing, and prepare a 10-15 minute presentation to class (there will be a concentrated presentations day end of semester).

- **Research-based**: investigate a research topic on your own (with instructor's guidance). For example, develop an algorithm or coding scheme and analyze it rigorously, extend a lower-bound technique shown in class, etc.).

- **Implementation-based**: There will be a (limited) selection of implementation projects, related to compression, coding and succinct data structures, which you can choose from. Depending on the project's scale, collaborations in small groups may be allowed.