Objectives of this course

- Teach some basic Information Theory;
- Illustrate (mathematical) power by applying it in combinatorics, complexity theory, algorithms, privacy, ...

- Such diversity? How can one person know it all & teach it? ... Err... I can’t + won’t.

- Course will be run seminar style. I will cover first few lectures, rest will be student run. We will learn together.

- Grades based on?
  - Projects
  - Presentations
  - Participation
  - Scribe work
Motivating Example: Shearer's Lemma

- Roughly relates volume of 3-d object to areas of its 2-d projections.

- Formally: Sets in finite universe; cardinalities & not volume/area.

Notation
- \([n] = \{1, 2, 3, \ldots, n\}\)
- for \(S = \{i_1, i_2, \ldots, i_k\} \subseteq [d]\)
- \(X = (x_1, x_2, \ldots, x_d) \in [n]^d\)
- \(X_S = (x_{i_1}, x_{i_2}, \ldots, x_{i_k})\)
- for \(F \subseteq [n]^d \land S \subseteq [d]\)
  \(F_S = \{x_S \mid x \in F\}\)
Shearer's Lemma \((k, d)\) - version

Let \( F \subseteq [n]^d \), then

\[
|F| \leq \left( \prod_{S \subseteq [d], |S| = k} |F_S| \right)
\]

\[
\Downarrow
\]

\(. \quad (k, 1) \) version:

\[
|F| \leq |F_{e_1}| \cdot |F_{e_2}| \cdot |F_{e_3}| \cdots |F_k|
\]

**Proof:** trivial

\( F \subseteq F_1 \times F_2 \times F_3 \times \cdots F_k \)

\(. \quad (3, 2) \) version

\[
|F|^2 \leq |F_{e_1}| \cdot |F_{e_2}| \cdot |F_{e_3}|
\]

Already non-trivial!
Proof via Information Theory

- Setup: - You & I know $F$.
  - I am given $w \in F$
  - I need to describe $w$ to you
  - How many bits do I need to send to you?

- Roughly $H(w)$ measures this quantity

- Notes: Entropy depends not only on $F$, but also how $w$ is chosen from $F$. 
Suppose further \( W = (x, y, z) \) & you and I already know \( z \), but Eve doesn't know \( z \).

- How many bits does Eve expect me to send you?

- \( H(W | z) \) denotes this quantity

\[ \text{Conditional Entropy of } W \text{ given } z \]

\[ \text{key player #2} \]

Aside: Today everything § definition, claims, proofs § will be handwavy. We will formalize everything later. But today we will see why such formalism may be useful.
Axioms Of (Conditional) Entropy

1. (What we should expect; and will later confirm)

2. \( w \in F \Rightarrow H(w) \leq \log_2 |F| \)

3. \( w \) uniform from \( F \Rightarrow |F| = 2^j \)

4. \( H(w) = j = \log_2 |F| \)

5. Suppose \( w = (x, y, z) \)

\[ H(x | y) \leq H(x) \]

(in the worst case, you and I can ignore \( y \) & focus on sending \( x \).

So \( x \) given \( y \) is easier to communicate.)

6. \( w = (x, y, z) \)

\[ H(x, y) = H(x) + H(y | x) \]

\[ = H(y) + H(x | y) \]
Remarkable fact:

- $H(x)$ satisfying above axiom exists!
- Will see in lectures 2, 3 ...

Return to Chapter

$\mathbf{2,1}$ - case:

$|F| \leq |F_1| \cdot |F_2|$

$\log |F| \leq \log |F_1| + \log |F_2|$

Pick $(x,y)$ uniformly from $F$

- $H(x,y) = \log |F|$
- $\log |F_1| \geq H(x)$ (not equal since projecting $(x,y) \to x$ in not uniform on $F_2$)
- $\log |F_2| \geq H(y)$
Suffices to show
\[ H(x, y) \leq H(x) + H(y) \]
\[ H(x, y) = H(x) + H(y|x) \]
\[ \leq H(x) + H(y) \]

\( (3,2) \) version of Shearer.

Wish to prove
\[ |F|^2 \leq |F_{12}| \cdot |F_{23}| \cdot |F_{31}| \]
\[ \Rightarrow \]
\[ 2H(x, y, z) \leq H(x, y) + H(y, z) + H(x, z) \]

\[ H(x, y) = H(x) + H(y|x) \]
\[ \text{RHS} \quad H(y, z) = H(y) + H(z|y) \]
\[ H(x, z) = H(x) + H(z|x) \]
\[ \text{LHS} \quad 2H(x, y, z) = 2H(x) + 2H(y|x) + 2H(z|x, y) \]

QED
Proof of $\phi_x(^1,k)$. Shearer

**Exercise** 😊

(Roughly: - Order variables
- Always condition on smaller #’ed variables
- Condition on all smaller #’ed variables to get LHS).

---

**Moral of the story**

- Somewhere in the axioms lies a very interesting inequality.
- Very useful!
- It has many other notions (Information, Divergence, HELLINGER distance)
- Many other inequalities (FANO, Pinsker, . . . . .)
- Many applications
In Computer Science

- Parallel Repetition
- Communication Complexity
  ↓
  Data Structure,
  Streaming algorithm
- Optimization
  "Max-Entropy Distribution"
- Crypto, Privacy, . . .
Punchline

I don't understand any of the above.

Goal of this course:

I should learn.

You will teach me. 😊

So what do I bring to the table

- many authors & course teachers are friends of mine

- I know good papers & can get answers to questions.
Way the course will work

- Next two lectures:
  I will review information theory basics

- Afterwards...
  - we will (jointly) select papers & present them.
  - Each lecture will have a paper + student covering it.
  - Student will read paper 1 week before & privately present to me. We will understand.
  - Andreina will skim paper & come in. We will discuss/present paper together.

- First paper
  "Entropy & Counting" - J.R.
General Info

Lecturer: Madhu Sudan
Office: MD 339
OH: 2:30-4 Thursdays
Email: madhu@cs.harvard.edu

Class Info

http://madhu.seas.harvard.edu/commex/Spring2016

Do join Piazza

To do items
- Sign up for scribing
- Paper + Date of Presentation
- Questions / Partner?
  Use Piazza
Come to OH