INFORMATION THEORY IN CS

Note Title 1/25/2016

Objectives of this course

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

———— x ————

- 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 $S = \{i_1, i_2, \ldots i_k\} \subseteq [d]$

  & $x = (x_1, x_2, \ldots x_d) \in [n]^d$

  $x_S \triangleq (x_{i_1}, x_{i_2}, \ldots x_d)$
- for $F \subseteq [n]^d$ & $S \subseteq [d]$

  $F_S \triangleq \{ x_S \mid x \in F \}$

Shearer's Lemma $\quad (k,d)$ - version

Let $F \subseteq [n]^d$, then

$$|F|^{\binom{d-1}{k-1}} \leq \left(\prod_{\substack{S \subseteq [d] \\ |S| = k}} |F_S| \right)$$

- $(k,1)$ version :

$$|F| \leq |F_{\{1\}}| \cdot |F_2| \cdot |F_3| \cdots |F_k|$$

Proof: trivial

$$F \subseteq F_1 \times F_2 \times F_3 \times \cdots F_k$$

- $(3,2)$ version

$$|F|^2 \leq |F_{12}| \cdot |F_{23}| \cdot |F_{31}|$$

Already non-trivial !

Proof via Information Theory

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

- Roughly $\underbrace{H(\omega)}$ measures this quantity

  $\boxed{\text{entropy of } \omega}$ ← key player!

- Notes: Entropy depends not only on F, but also how ω 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

$\downarrow$

Conditional Entropy of w given Z

key player #2

<u>Aside</u>: Today everything {definitions, 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

- (what we should expect; and will later confirm)
- $W \in F \Rightarrow H(W) \le \log_2 |F|$
- W uniform from F & $|F| = 2^j$

  $\Rightarrow H(W) = j = \log_2 |F|$

  will ignore

- Suppose $W = (X, Y, Z)$

  $H(X|Y) \le H(X)$

  (in the worst case, you and I can ignore Y & focus on sending X. So X given Y is easier to communicate.)

- $W = (X, Y, Z)$

$$H(X,Y) = H(X) + H(Y|X)$$
$$= H(Y) + H(X|Y)$$

Remarkable fact :

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

—— x ——

Return to Shearer

$(2,1)$ - 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 $(X,Y) \to X$ is not uniform on F_1)
- $\log |F_2| \geq H(Y)$

Suffices to show

$$H(X,Y) \le H(X) + H(Y)$$

$$H(X,Y) = H(X) + H(Y|X)$$
$$\le H(X) + H(Y) \qquad \boxtimes$$

(3,2) version of Shearer

wish to prove

$$|F|^2 \le |F_{12}| \cdot |F_{23}| \cdot |F_{31}|$$
$$\Uparrow$$
$$2H(X,Y,Z) \le H(X,Y) + H(Y,Z) + H(X,Z)$$

$$H(X,Y) = H(X) + H(Y|X)$$

RHS $$H(Y,Z) = H(Y) + H(Z|Y)$$

$$H(X,Z) = H(X) + H(Z|X)$$

$$\qquad \geq \qquad \geq \qquad \geq$$

LHS $$2H(X,Y,Z) = 2H(X) + 2H(Y|X) + 2H(Z|X,Y)$$

$\boxed{\text{QED}}$

Proof of (k,d)-Shearer

$\boxed{\text{Exercise}}$ ☺

(Roughly : - order variables
- always condition on smaller #'ed variables
- Condition on <u>all</u> smaller #'ed variables to get LHS).

<u>Moral of the story</u>

- Somewhere in the axioms lies a very interesting inequality.
- Very^n useful !
- IT has many other notions (Information, Divergence, Hellinger distance)
- Many other inequalities (Fano, Pinsker, ...)
- many applications

In Computer Science

- Parallel Repetition
- Communication Complexity

  $\Downarrow$

  Data Structure,

  Streaming algorithms
- Optimization

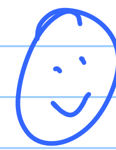
  "Max-Entropy Distribution"
- Crypto, Privacy, ...

... 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/courses/Spring2016

Do Join Piazza

To do items

- Sign up for scribing
- Paper + Date of Presentation.
- Questions / Partner ?

  Use Piazza

  Come to OH.