CS286r: Topics at the Interface between Computer Science and Economics

Fall 2010

Information, Prediction, and Collective Intelligence

Yiling Chen
yiling@eecs.harvard.edu

Course website: http://www.eecs.harvard.edu/cs286r/

Computer Science

- is the study and the science of the theoretical foundations of information and computation and their implementation and application in computer systems. [Wikipedia]
  - Building systems
  - F(x) =?
  - How fast can we get the answer?
  - Focus on computational and informational constrains.

Economics

- is the social science that studies the production, distribution, and consumption of goods and services. [Wikipedia]
  - Economies (systems)
  - Many self-interested agents
  - Agents' preferences/utilities over outcomes
  - Agents' information and beliefs
  - Agents' decision making
  - Game-theoretic interactions of agents

The Interface

- Computer systems are increasingly being developed and used by multiple parties with different preferences
  - Predict system outcomes
  - Design systems to achieve desired outcomes
- Economic problems sometimes are (hard) computational problems
  - Resource allocation
  - Price discovery

Theories, algorithms, and systems that satisfy both economic and computational constraints.

Lots of Compelling Applications

- Internet Monetization:
  Google, Yahoo!, Microsoft are using auctions to sell ads
- Markets are used for information aggregation
  - Google, Yahoo!, Microsoft, GE, etc. have internal prediction markets
- Social network and Social Tagging:
  Facebook, MySpace, LinkedIn, Flickr, LibraryThing

…

This Course

- Rotating topic course
- Previous
  - Fall 2009. Assignment, Matching, and Dynamics
  - Fall 2008. Social Computing
  - Spring 2008. Computational Finance
  - Spring 2007. Computational Mechanism Design
  - Spring 2006. Multi-agent Learning and Implementation

  …
- Seminar style

Course Goals

- Provide an introduction to an emerging, interdisciplinary literature
- Develop a level of comfort with both economic and computational thinking
- Develop general skills related to reading papers, identifying research questions
- Provide a basis for continued research.

Fall 2010

- Information, Prediction, and Collective Intelligence
- Algorithmic, game theoretic, and conceptual questions related to obtaining information, making predictions, and getting tasks done by the crowds.
- Focus on *eliciting and aggregating probabilistic information*, bridging from economic theory to theory of online learning.

Crowds Are Smarter…

- Who wants to be a millionaire?
  - Fifty-Fifty
    Correct 50% of the time
  - Phone-A-Friend
    Correct 65% of the time
  - Ask the Audience
    Correct 91% of the time

Crowds Are Smarter…

- Jelly-Beans-in-the-Jar Experiment
  - Professor Jack Treynor ran the experiment in his class
  - with a jar that held 850 beans
  - the group estimate was 871
  - only one of the 56 people in the class made a better guess

Are Crowds Smarter?

- No always
  - Bad committee decisions
  - Endless group meetings
- In this course, we focus on mechanisms that intend to make crowds smarter.

Structure of the Course

- Introductory lectures (5 lectures)
  - This one, information theory and Kelly criterion, and game theory
- Research Papers
  - Incentivizing experts
  - Peer prediction
  - Prediction markets
  - Online learning
  - Analysis of existing collective intelligence systems

Enrollment & Prerequisites

- Enrollment is limited to about 20 students. Complete Survey at end of class!
- Prerequisites
  - Math background is important! At least a basic course in linear algebra (such as M 21b, AM 21b, or equivalent)
  - A course on probabilities and statistics (STAT 110 or equivalent)
  - An algorithm course (CS 124, or equivalent)
  - Familiarity with the concept of rationality. An AI course or an economics/game theory course.

Advanced course in algorithms, microeconomics, game theory, or linear programming are helpful but **not** required.

Grading

Problem sets	25%	2-3 homework problem sets
Participation	25%	Reading papers, submitting short summaries and questions before class, and participation in class discussion. (Note: Absent students rarely contribute to discussions.)
Presentation of one or two research papers	15%	A short survey and critique of the papers. Lead class discussion.
Project	35%	Project proposal, class presentation, and final report.

Project

- **Goal:** develop a deep understanding of a specific research area and to the extend possible to work on an open research problem.
- Can be theoretical, computational, experimental, or empirical.
- Can write an exposition paper, but needs novelty!
- A list of high-level project topics will be provided. **You are encouraged to propose your own topic for approval!**
- Tentative project due dates:
  - Wednesday 11/3: project proposal due
  - Monday 12/6: brief project presentation
  - Friday 12/10: project report due

Logistics

- TF
  - Alice Gao
- Office Hours
  - Yiling: Monday 2:30 – 3:30, MD 339
  - Later will add office hours likely on Thursdays to meet with students in advance of presenting papers
  - Today, 2:30 – 3:30
  - Alice: Tue 2:30-4, MD 242

Missed course materials from the TF

Information, Prediction, and Collective Intelligence

- Motivating examples
  - Disease surveillance
  - Business forecasting
  - Multi-agent systems
  - Almost any decision making under uncertainty
  - Netflix Prize
  - DARPA Red Balloon Challenge

Examples of Course Topics

- Incentivizing experts
- Peer prediction
- Prediction Markets
- Online learning
-

Incentivize Experts

- Suppose I'd like to get information about tomorrow's weather (sunny or rainy?)
- How can I ensure that an expert will tell me his/her **true** probability assessment of the event?

What If We Won't Know the Outcome?

- Eg. Conditional events, subjective information
- Surveys
  - Eg. How many hours per week you spent on assignments?
    - Less than 5 hours
    - 5-10 hours
    - 10-20 hours
    - Above 20 hours

Peer Prediction and Bayesian Truth Serum

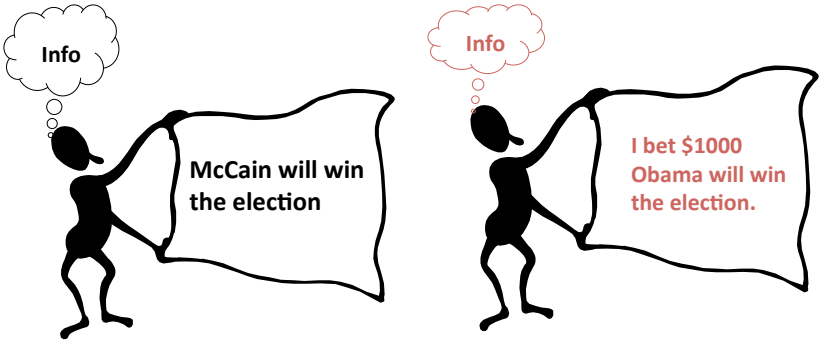
Combining information is hard!

- If we have multiple experts, how can we combine their information?
- Some impossibility results on combining probability distributions.
  - T(f1, f2, f3, …, fn)
  - External Bayesianity
  - Independent of irrelevant alternatives

  => dictatorship

Bet = Credible Opinion

- Q: Will Obama win the Presidential election?



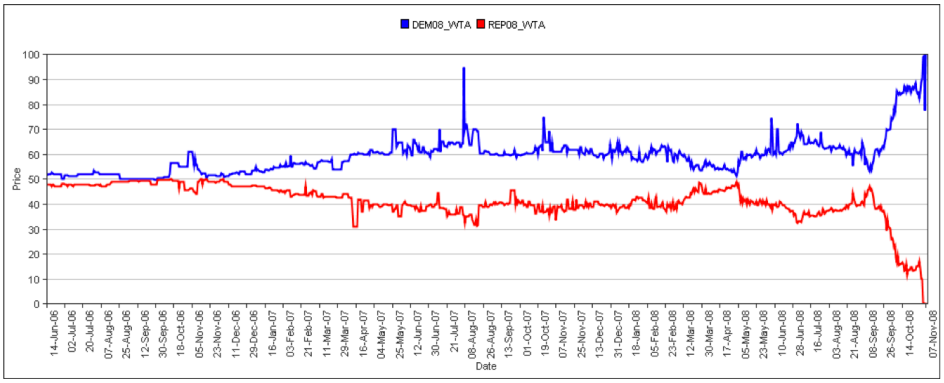
- Betting intermediaries
  - Las Vegas, Wall Street, Betfair, Intrade,...

Prediction Markets

- A prediction market is a futures market (betting intermediary) that is designed for information aggregation and prediction.
- Payoffs of the traded item is associated with outcomes of future events.

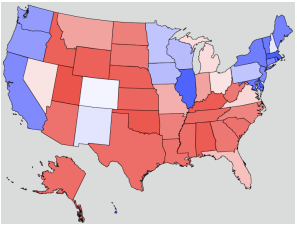
$f(x)

Example: Iowa Electronic Market



A Combinatorial Betting Example

- 2^{51} outcomes, $2^{2^{51}}$ combinations
- Allow participants to bet on logical formulas
  - Create contracts on the fly:
    $1 if Ohio AND Florida OR New York, $0 otherwise
  - Specify buy price and quantity
- Computationally hard!



Learning from Expert Advice

Learning from Expert Advice

- The algorithm maintains weights over N experts

Slide Source: J.W. Vaughan

Learning from Expert Advice

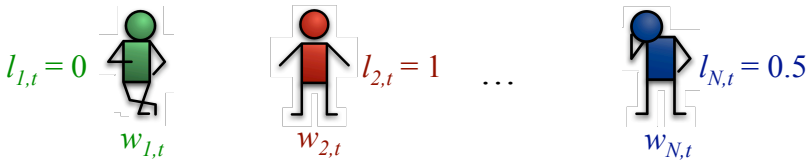
- The algorithm maintains weights over N experts

- At each time step t, the algorithm…
  - Observes the instantaneous loss $l_{i,t}$ of each expert i

Slide Source: J.W. Vaughan

Learning from Expert Advice

- The algorithm maintains weights over N experts

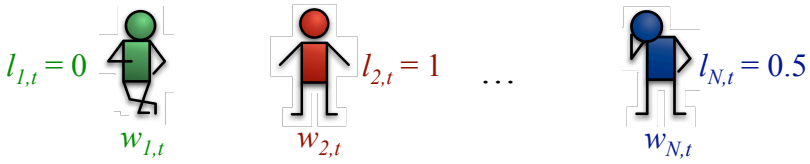


- At each time step t, the algorithm…
  - Observes the instantaneous loss $l_{i,t}$ of each expert i

Slide Source: J.W. Vaughan

Learning from Expert Advice

- The algorithm maintains weights over N experts

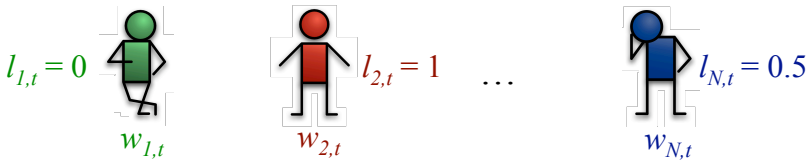


- At each time step t, the algorithm…
  - Observes the instantaneous loss $l_{i,t}$ of each expert i
  - Receives its own instantaneous loss $l_{A,t} = \Sigma_i \, w_{i,t} \, l_{i,t}$

Slide Source: J.W. Vaughan

Learning from Expert Advice

- The algorithm maintains weights over N experts

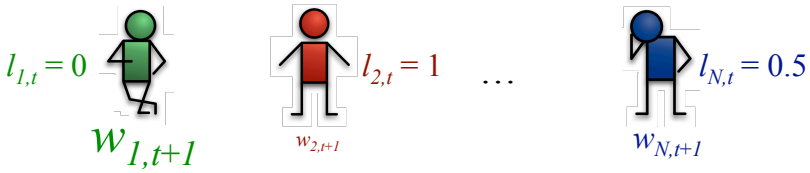


- At each time step t, the algorithm…
  - Observes the instantaneous loss $l_{i,t}$ of each expert i
  - Receives its own instantaneous loss $l_{A,t} = \Sigma_i\, w_{i,t}\, l_{i,t}$
  - Selects new weights $w_{i,t+1}$

Slide Source: J.W. Vaughan

Learning from Expert Advice

- The algorithm maintains weights over N experts



- At each time step t, the algorithm…
  - Observes the instantaneous loss $l_{i,t}$ of each expert i
  - Receives its own instantaneous loss $l_{A,t} = \Sigma_i\, w_{i,t}\, l_{i,t}$
  - Selects new weights $w_{i,t+1}$

Slide Source: J.W. Vaughan

Learning from Expert Advice

- **Classic result:** There exist algorithms such that on any (bounded) sequence of losses,

$$\text{algorithm's cumulative loss} - \text{loss of the best performing expert} < O((T \log N)^{1/2})$$

Slide Source: J.W. Vaughan

Learning from Expert Advice

- **Classic result:** There exist algorithms such that on any (bounded) sequence of losses,

$$\text{algorithm's cumulative loss} - \text{loss of the best performing expert} < O((T \log N)^{1/2})$$

- Holds even in a fully adversarial setting with no statistical assumptions about the sequence of losses

Slide Source: J.W. Vaughan

Learning from Expert Advice

- **Classic result:** There exist algorithms such that on any (bounded) sequence of losses,

$$\underbrace{\text{algorithm's cumulative loss} - \text{loss of the best performing expert}}_{\text{“regret”}} < O((T \log N)^{1/2})$$

- Holds even in a fully adversarial setting with no statistical assumptions about the sequence of losses

Slide Source: J.W. Vaughan

Learning from Expert Advice

- **Classic result:** There exist algorithms such that on any (bounded) sequence of losses,

$$\underbrace{\text{algorithm's cumulative loss} - \text{loss of the best performing expert}}_{\text{“regret”}} < O((T \log N)^{1/2})$$

- Holds even in a fully adversarial setting with no statistical assumptions about the sequence of losses

- Algorithm is said to have “no regret” since the average regret per trial approaches 0 as T grows

Slide Source: J.W. Vaughan

For Wed. 9/8

- Submit comments on Chapter 2 (2.1 – 2.8) of Elements of Information Theory
- Reading is posted on the class schedule.
- Check course website later this week on how to submit your comments
- What is unclear? What would you like to hear about in class? What did you enjoy?

- Please hand in the survey now.