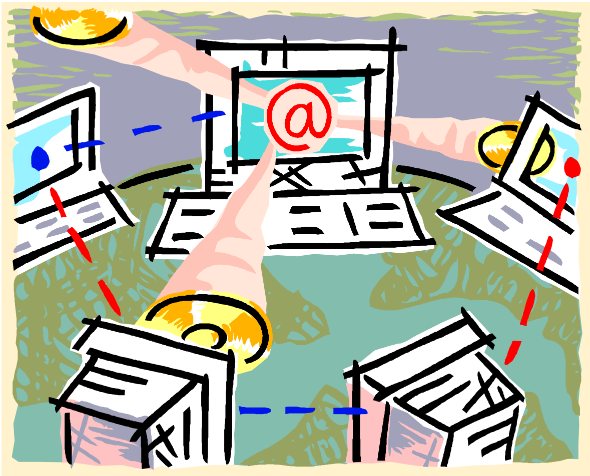
Social Choice and Networks

Elchanan Mossel
UC Berkeley



Logistics 1

- Different numbers for the course:
- Compsci 294 Section 063
- Econ 207A
- Math C223A
- Stat 206A
- Room: Cory 241
- Time TuTh 17-18:30
- (is a preference for 17-20 on one of the days)?

Logistics 2

- If we stay in large #s we will need a GSR

- If you want to be the GSR – send me an email with your c.v. (will secure a sit in class ☺)

- I will look into moving into a bigger class.

- There will be a link to course materials from my home page:

- stat.berkeley.edu/~mossel

Prerequisites

The course is aimed at graduate students in Statistics, EECS, Mathematics, Economics etc.
It assumes working knowledge of probability, discrete mathematics and computer science at a graduate level. In particular, the following will be assumed (among many other topics):

1. Probability: conditional expectations, martingales and multi-variate Gaussian variables.

2. Linear algebra including eigenvalues etc.

3. Understanding of algorithms and computational complexity analysis.

4. Basic graph theory.

What do I need to get a grade?

1. Each week: Group polish presentation in ppt,
   write it as a book chapter.
   generate HW problems from the lecture.

2. HW sets – in groups. These will be hard about every 2 weeks.

3. Final exam (in class or take home to be decided)

4. + -> projects
5. Networking aspect: 2-4 people in a group – cannot be all from the same department.
6. Grade = 0.4*(exam) + 0.3*(lecture) + 0.3*(HW) + projects

The “Big Questions”

- How are collective decisions made by:
- people / computational agents

- Examples: voting, pricing, measuring.

- Biased / Unbiased private signals.

- Network structure.

- Types of signals (numbers, binary, behaviors etc.)

- Opinion leaders, communities.

Topic 1: Aggregation of Biased Signals

Condorcet's jury theorem (*Essay on the Application of Analysis to the Probability of Majority Decisions*, 1785):

• a group wishes to reach a decision by majority vote.

•One of the two outcomes of the vote is *correct*, and

• each voter votes correctly independently with probability $p>1/2$.

•Then in the limit as the group size goes to infinity, the probability that the majority vote is correct approaches 1.

Topic 1: Aggregation of Biased Signals

In this course:

- A proof of Condorcet Theorem & Extensions:
- What are the best/worst functions for aggregation (e.g. electoral college, dictator etc.)?
- Condorcet's theorem for more than 2 alternatives.
- An aggregation theorem for non independent voters.

Topic 1: Aggregation of Biased Signals

In this course:

- Algorithmic problem:
- How to aggregate rankings of many alternatives all “correlated” with truth.
- Example: Search engines.

Topic 2: Agreeing to Disagree and Repeated Voting

- Basic questions: Can people agree to disagree? What can be gained from discussions?
- Example: A good student makes good impression in a class with probability 90%. A bad student makes good impression with probability 70%.
- Given 50 professors what is a good way to do decide who is a good student and who is not?
- In this class: Aumann’s agreeing to disagree in the Bayesian setup
- A computationally efficient version.

Topic 3: Biased Signals on Social Networks

- Models for repeated voting on social networks.
- Only see your neighbors votes.
- Basic questions: how well is information aggregated? Do all voters converge to same opinion/vote?

Topic 3: Biased Signals on Social Networks

- Heuristic models for repeated estimation on social networks:

- DeGroot, Jackson: Repeated averaging with neighbors.Easiest to analyze and define social influence etc.

- Majority dynamics (easy to define, hard to analyze).
- Computational framework and social experiments.

Topic 3: Biased Signals on Social Networks

- Bayesian models:
- Complete Bayesian models.
- Computational aspects.

Topic 4:Unbiased / conflicting signals

- No good decision rules:

- Arrow theorem -> no rational way to rank
- GS theorem -> no non manipulable way to elect a winner.

- Clustering problems:
- communities based on connectivity.
- How to define communities based on behaviors and interactions.

Topic 4:Unbiased / conflicting signals

- The algorithmic problem of finding the best set to market to.
- Hardness.
- Submodular case.
- The bootstrap percolation case.
- Probabilistic models: sharp thresholds and competition models on networks.

Topic 5:Other aspects of network study

- Network models and power laws.

- Data Mining and Search on Networks.

Check your Math quiz

- Take a piece of paper out.
- Write your name and student i.d. on it.
- The following <u>will not</u> be part of your grade
- But:
- A necessary but very insufficient condition to be able to pass this course is to easily answer all of the questions below in < 10 mins.

Check your Math quiz

Which of the following questions can you answer?

1. Show that for any finite graph the sum of the degrees of all nodes is even.
2. Let (X,Y) be a normal vector with E[X] = E[Y] = 0
   $E[X^2] = E[Y^2] = 1$ and E[X Y] = 0.5.
   What is E[X|Y]?
3. eigenvalues and singular values of A = [0,1 ; 0,1] ?
4. How many comparisons are needed to sort n elements? Why is the lower bound true?

Check your Math quiz

1. Show that for any finite graph the sum of the degrees of all nodes is even.

A. This follows from the fact that each edge contributes 2 to the sum of the degrees.

2. Let (X,Y) be a normal vector with $E[X] = E[Y] = 0$
   $E[X^2] = E[Y^2] = 1$ and $E[X\ Y] = 0.5$.
   What is $E[X|Y]$?

A. $X = 0.5Y + Z$ where Z is normal of mean 0 so
   $E[X \mid Y] = 0.5Y$.

Check your Math Solutions

3. eigenvalues and singular values of A = [0,1 ; 0,1] ?

A. The matrix is diagonal so the eigenvalues are 0,1.

   For the singular value we need to find the eigenvalues of A A* = [1, 1; 1,1] with eigenvalues 0 and 2. The singular values are 0,2^{1/2}.

4. How many comparisons are needed to sort n elements? Why is the lower bound true?

- A Order n log n. The lower bound follows since there are n! permutations and 2^k possible outputs where k is the number of comparisons.