Behavioral Network Science: Individual Modeling and Collective Behavior

Michael Kearns
Computer and Information Science
University of Pennsylvania

First Year Review, August 27, 2009
Kearns
Behavioral network science

Theory
• First principles
• Rigorous math
• Algorithms
• Proofs

Data Analysis
• Correct statistics
• Only as good as underlying data

Numerical Experiments
• Simulation
• Synthetic, clean data

Lab Experiments
• Stylized
• Controlled
• Clean, real-world data

Field Exercises
• Semi-Controlled
• Messy, real-world data

Real-World Operations
• Unpredictable
• After action reports in lieu of data
Network Science in the Kearns Group: Ongoing and Recent Projects

• Behavioral Experiments in Network Science
  – S. Judd, E. Vorobeychik, T. Chakraborty, J. Tan, J. Wortman, MK
• Network Structure and Strategic Outcomes
  – T. Chakraborty, J. Tan, MK
• Local Algorithms for Global Network Properties
  – M. Brautbar, MK
• Theory of Learning from Collective Behavior
  – J. Wortman, MK
Experimental Agenda

- Human-subject experiments in collective and individual strategic behavior in networks
- Subjects simultaneously participate in groups of ~ 36 people
- Subjects sit at networked workstations
- Each subject controls some simple property of a single vertex in some underlying network
- Subjects have only *local* views of the activity: state of their own and neighboring vertices
- Subjects have (real) financial incentive to solve their “piece” of a collective (global) problem
- Across many experiments, have deliberately varied *network structure* and *problem/game*
  - networks: inspired by models from network science (small worlds, preferential attachment, etc.)
  - problems: chosen for diversity (cooperative v. competitive) and (centralized) computational difficulty
Experiments to Date

- **Graph Coloring (Jan 2006; Feb 2007)**
  - player controls: color of vertex; number of choices = chromatic number
  - payoffs: $2 if different color from all neighbors, else 0
  - max welfare states: optimal colorings
  - centralized computation: hard even if approximations are allowed

- **Consensus (Feb 2007)**
  - player controls: color of vertex from 9 choices
  - payoffs: $2 if same color as all neighbors, else 0
  - max welfare states: global consensus of color
  - centralized computation: trivial

- **Independent Set (Mar 2007)**
  - player controls: decision to be a “King” or a “Pawn”; variant with King side payments allowed
  - payoffs: $1/minute for Solo King; $0.50/minute for Pawn; 0 for Conflicted King; continuous accumulation
  - max welfare states: maximum independent sets
  - centralized computation: hard even if approximations are allowed

- **Exchange Economy (Apr 2007)**
  - player controls: limit orders offering to exchange goods
  - payoffs: proportional to the amount of the other good obtained
  - max welfare states: market clearing equilibrium
  - centralized computation: at the limit of tractability (LP used as a subroutine)

- **Biased Voting (“Democratic Primary Game”; May 2008)**
  - player controls: choice of one of two colors
  - payoffs: only under global agreement; different players prefer different colors
  - max welfare states: all red and all blue
  - centralized computation: trivial

- **Networked Bargaining (May 2009)…**
Biased Voting: Quick Review

[ MK, S. Judd, J. Tan, J. Wortman; PNAS Jan 2009 ]
Democratic Primary Games

Zak Xavier

If unanimity is reached, your payoff will be:

$0.75 for red, $1.25 for blue
Comparison of capped action changes (blue = coER, green = coPA, blue = power)

Mean number of color changes in behavioral experiments

Mean number of action changes in simulations, each trial capped at 250

[MK, S. Judd, J. Tan, J. Wortman; PNAS Jan 2009]
Individual and Collective Modeling in Coloring and Consensus
[S. Judd, E. Vorobeychik, MK]
game progress: 80%

your current payoff: $2.00

(gameoff is $2.00 if your color is DIFFERENT from all your neighbours, otherwise $0.00)
Network Formation Model

• Single parameter $p$ (a probability)
• $p=0$: a chain of 6 cliques of size 6 each (see figure)
• $p > 0$: each intra-clique edge is “rewired” with probability $p$:
  – first flip $p$-biased coin to decide whether to rewire
  – if rewiring, choose one of the endpoints to “keep” the edge
  – then choose new random destination vertex from entire graph
• Values of $p$ used: 0, 0.1, 0.2, 0.4, 0.6 1.0
• Three trials for each value of $p$; different random network for each trial
• Move from “clan” topology to random graphs
Network Properties vs. Rewiring Probability
Summary of Events

- Held 18 coloring and 18 consensus experiments
- All 18 coloring experiments globally solved
  - average duration ~ 35 seconds
- 17/18 consensus experiments globally solved
  - average duration ~ 62 seconds
- Recall (worst-case) status of problems for centralized computation
  - seems it is easier to get people to disagree than to agree…
Influence of Structure

large p: easier for consensus
large p: harder for coloring (cf. Jan 2006 experiments)
small-p vs large-p performance statistically significant
Styles of Play: Meaningful Individual Variation

• “Random Observer” methodology
• Frequency of color changes
• Time of first entry
• Fraction of “suboptimal” play

CONSENSUS

COLORING

First entry
Frequency of change
Suboptimal proportion
Behavioral Modeling

- Assume players make decisions in discrete (1s) time steps
- Two-part model:
  - activation: player becomes active with some probability
  - decision: if a player is active, she chooses the “best” color (most/least neighbors picking this color in consensus/coloring)
- Decision model is fixed
- Activation model is learned from experimental data
  - activation probability: logistic regression
  - what worked: single feature, proportion of neighbors playing my previously chosen color; learned from data for all subjects
  - what didn’t:
    - subsets of other features: degree, progress, indicator whether last color is “suboptimal” (a better one can be played), number of rounds last color has been suboptimal, proportion of game played
    - a variety of alternative logistic regression models to determine the color choice
    - individual-level models (not enough data?)
average experiment duration vs p: coloring (blue), consensus (red)

consensus

coloring

ONR MURI: NexGeNetSci
Art by Consensus
(small p)