Motivation

- Lots of human and business activity is about **coordinated decision making**
  - how many of your decisions are made in isolation?
- **Private information** and **self-interest** is intrinsic to many of these environments.
- Useful artificial intelligence must meet this demand: accelerate, simplify and improve group decision making and problem solving.
- Lots of network problems as well:
  - bandwidth allocation (“smart networks”)
  - grid computer resource allocation

Example Scenarios I

- Meeting scheduling:
  - multiple parties, preferences and constraints
  - design interaction protocol to enable good schedules
- Grid computing:
  - multiple parties, preferences and constraints
  - use of grid computation to run data- and compute-intensive experiments
- E-commerce:
  - markets (e.g., eBay) to facilitate trade
  - sponsored-search auctions
  - industrial procurement

Example Scenarios II

- Task allocation:
  - allocate tasks to teams (e.g., Crimson editors)
  - respect ordering constraints, preferences, etc.
- Sequential decision making
  - e.g., a software deciding each week how to set priorities across different projects
  - uncertainty, private information, etc.
- “Market of minds”
  - coordinated learning for problem solving
  - how to provide incentives s.t. self-interested agents learn to specialize, work effectively as a team
Shared Attributes

- Multiple parties
- Private information and self-interest
- Uncertainty
- Repeated, or sequential
- Design goals

⇒ multi-agent learning and implementation

Bridging Two Communities

- **Economics**
  - traditional emphasis on game-theoretic rationality, describing, and systems (“economies”) with *multiple* self-interested agents.
- **Computer Science**
  - traditional emphasis on computational & informational constraints, *building*, and systems with one user (or a user with adversarial opponents)

CS286r: Unification, resolve conflicts between game-theoretic and computational constraints, develop rigorous theories, methodologies and algorithms.

Very Active Research Area

- **Seventh ACM Electronic Commerce Conference**, June 2006 (Ann Arbor MI)
- **Dagstuhl Workshop on Electronic Market Design**, Jan 2005 (Dagstuhl, Germany)
- **Games & Theoretical Economics (GATE)** Workshop, Northwestern Univ Nov 2005 (Evanston, IL)
- **WINE 2005: The 1st Workshop on Internet and Network Economics**, Dec 2005, Hong Kong
- **Second Bertinoro Workshop on Algorithmic Game Theory (AGATE)** 2-6 July 2006, Forli Italy

MAL and Implementation

- **Why Learning?**
  - coordinated decision making is often *sequential*, or *repeated*
  - adaptive behavior as bounded-rationality
  - environments are often uncertain, learning can improve performance
- **What is “Implementation”?**
  - term from game theory
  - loosely, *making good joint decisions in multi-agent systems despite private information and self-interest*
  - provides theories on how to design incentives
- **Together**: design to promote good outcomes in systems with adaptive and self-interested agents.
A Bird’s Eye View

- **Mechanism design and Implementation theory:** in its focus on designing incentives, rules-of-encounters
- **Reinforcement Learning:** well-understood for single-agent domains, being extended to multi-agent domains
- **Game theory:** learning as a way to converge on an equilibrium in repeated games (often to “explain” NE)
- **Distributed AI:** design to promote good, coordinated decision making (w/ collaborative agents)

Mechanism Design, Implementation theory

- “Inverse game theory.” Design rules to promote good outcomes in equilibrium.
- Mechanism design:
  - trusted center; receives reports from agents;
  - makes and enforces decisions.
- game of incomplete information
- Implementation theory:
  - (more) dis-intermediated; self-policing by agents
  - complete information about types
  - group of agents self-enforce outcomes

Reinforcement Learning

- Single-agent Markov Dec. Process (MDP): $s \in S, a \in A, r(s, a) \geq 0, Pr(s, a, s') \in [0, 1]$.
- **Q-learning:** Maintain $Q(s, a)$ values. Repeat. Choose $a_t$ from $s_t$ (e-greedy). Take action $a_t$. Observe $r(s_t, a_t), s_{t+1}$. Update: $Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha[r(s_t, a_t) + \gamma V(s_{t+1})]$. Here, $V(s) = \max_a Q(s, a)$.
- Markov Game: $(N, S, A = [A_1, \ldots, A_N], r = [r_1, \ldots, r_N], Pr(s, \bar{a}, s'))$. Reward to agent $i$, $r_i(s, \bar{a})$, depends on joint actions.
- How to generalize single-agent RL to multi-agent RL? What should the goal be? (Self-play? Optimality? Equil.?)

Multi-Agent RL

- Pretend environment is passive. Each agent does standard Q-value update.
- Define Q-values as function of all agents’ actions. $Q_i(s_t, \bar{a}_t) \leftarrow (1 - \alpha)Q_i(s_t, \bar{a}_t) + \alpha[r_i(s_t, \bar{a}_t) + \gamma V_i(s_{t+1})]$
- But, how to define the $V_i$ function? (Intuitively, should represent the sum of Q-values for max joint action.)
- Various proposals: minimax-Q, belief-learning, Nash-Q, CE-Q, ...
- Unlike single-agent learning, limited convergence results.
Game-theory and Learning

- Repeated games (special case of Markov games).
  Study convergence to equilibrium play. (to what?)
- Belief learning. E.g., fictitious play. Maintain a model to predict play of opponent. Play BR to model.
- Bayesian learning. E.g., Kalai and Lehrer.
- Simple Behavioral rules. E.g., Erev and Roth.
- No-regret based approaches. E.g., Fudenberg and Levine.
- Rich possibility and impossibility results (e.g., Nachbar; learnability, richness, consistency...)

Distributed AI

- Heuristic (somewhat ad hoc) design of mediated, and partially-mediated multi-agent systems
- Hierarchical RL: a center assigns reinforcements to individual agents based on team reinforcement
- Social rules: constraints on actions of agents to promote better outcomes
- Collectives: design of informative signals to promote fast convergence on good joint policies.

Course Goals

- **Main goal:** find a synthesis, towards “learned implementation.” (Promote fast, tractable learning, in equilibrium, by self-interested agents towards desirable outcomes.)
- **Secondary goals:**
  - categorize complex game-theory literature on learning (info. assumptions, guarantees, poss & impos, etc.)
  - categorize complex AI literature on MAL (understanding objectives, info. assumptions, guarantees, critiques, etc.)
  - familiarity w/ “implementation”, from both an economics and an AI perspective
  - read and get to know some fun papers!

What is this Course?

**Rotating topics course.**

- Spring, 2006. Multi-agent Learning and Implementation

**Tentative future plans:**

- Spring, 2007. Computational Mechanism Design
- Spring, 2008. Not planned to offer the class.

**Previous years:**

- Spring 2003: Electronic market design
- Spring 2004: Iterative combinatorial exchanges
**Course Structure**

- **Introductory lectures:** [until 2/22]
  - (5 in total) game theory, mechanism design, Nash implementation, sequential decision making & reinforcement learning
- **Current Research papers:** [bulk of the course]
  - Learning in Repeated Games; Markov Games
  - Multi-agent RL; Coordinated RL
  - Design of Collectives; Implementation theory
  - Partial-control of Multiagent systems

**Prerequisites, Enrollment**

Enrollment limited to around 20 students. Complete survey at end of class!

**Ambitious class:** attempting a new synthesis.

**Technical papers,**

- Level of math. sophist., at least a basic course in linear algebra (such as M 21b, AM 21b, or equivalent)
- Theoretical CS, for example CS 121 (complexity theory) and CS 124 (algorithms).
- Familiarity with “agent rationality” concepts, e.g., an AI course, CS 181 or CS 182, or an appropriate econ/GT course.

Advanced algorithms, game theory, microeconomics, linear programming, all very helpful (not required).

**Grading**

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem sets</td>
<td>25%</td>
</tr>
<tr>
<td>Participation</td>
<td>20%</td>
</tr>
<tr>
<td>Presentation of</td>
<td>15%</td>
</tr>
<tr>
<td>Research paper</td>
<td></td>
</tr>
<tr>
<td>Project</td>
<td>40%</td>
</tr>
</tbody>
</table>

- chat with me to place out

**Office Hours**

Regular hours:
Tues Thurs 2-3pm, Maxwell-Dworkin 229.

Extra early-term office hours: Wed 2/1, 2:30-3pm; Mon 2/6 2:30-3pm.

**Teaching Staff:**
Ben Lubin, MD 205, blubin@eecs.harvard.edu
Giro Cavallo, MD 209, cavallo@eecs.harvard.edu

oh’s TBD, (will hold sections, starting next week.)

Come talk to us about your background for the course.
Projects

- **Goal:** Develop a deep understanding of an important open research area, and to the extent possible, to work on an open research problem.
- Projects may be theoretical or experimental. (Experimental projects in groups of two.)
- Suggested topics provided, although can submit own subject to approval.
- Proposals due after Spring Break; Initial presentations on the last day of class; Project reports before Finals.

Example Projects from 2005 (CMD)

- Collusion vs. False-name Proofness in CAs
- A study on privacy-preserving Proofness
- Continuous combinatorial exchanges w/ applications to Course-allocation Aftermarkets
- An Empirical Study of the Lavi-Nisan online auction
- Inefficiencies in selfish routing

Introductory Reading

- On Learnable MD (Parkes)
- On the Agenda of Research in Multi-agent Learning (Shoham et al.)
- Rationality and Bounded rationality (Aumann)

I: On Learnable MD

- Choose from $K$ outcomes each period. Let $f(v) = \arg\max_{k \in K} \sum_i v_i(k)$. Values $v_i(k) \geq 0$. Same agents $v = (v_1, \ldots, v_N)$ each period.
- VCG mechanism:
  1. report $\bar{\theta} = (\bar{\theta}_1, \ldots, \bar{\theta}_N)$
  2. choose $\hat{k}^*(\hat{\theta}) = f(\hat{\theta})$
  3. pay $p^*(\bar{\theta}) = \sum_{j \neq i} \bar{\theta}_j(\hat{k}^*(\bar{\theta})) - h_{-i}(\bar{\theta})$
- Agent $i$'s utility:
  $u_i(\bar{\theta}) = v_i(\hat{k}^*(\bar{\theta})) + \sum_{j \neq i} \bar{\theta}_j(\hat{k}^*(\bar{\theta})) - h_{-i}(\bar{\theta})$
- **Goal:** design $h_i$ to promote fast learning of the equilibrium; and thus implement $f(v)$.
**COlleCtive INtelligence (COIN)**

(Tumer and Wolpert)

Goal: facilitate collectives via informative feedback

- $u_i(\mathbf{v}) = u_i(k^*(\mathbf{v})) + \sum_{j \neq i} \hat{v}_j(k^*(\mathbf{v})) - h_i(\mathbf{v})$

- (Groves) $h_i(\mathbf{v}) = 0$

- (VCG) $h_i(\mathbf{v}) = \hat{v}_i(k^*(\mathbf{v})) + \sum_{j \neq i} \hat{v}_j(k^*(\mathbf{v}))$

- (WLU) $h_i(\mathbf{v}) = \sum_{j \neq i} \hat{v}_j(k^*(\mathbf{v})) - \max_k \sum_{j \neq i} \hat{v}_j(k)$

- E.g., single-item auction:
  - *win*, pay $\max(\tau, v(2))$
  - *lose*, pay $\max(0, \tau - v(1))$

where $v(1), v(2)$ are the max, and 2nd-max.

---

**Example: Auction Game**

- Simple learning (Boltzmann exploration)

- Agent $i$ bids $\hat{v}_i = kv_i$, $k \in \{0.6, \ldots, 1.1\}$.

  Maintains estimate $\pi^t_i(k) \geq 0$ for each $k$. Plays $k$ prop. to $e^{\pi^t_i(k)} / \text{temp}^t$, where $\text{temp}^{t+1} = a \times \text{temp}^t$, $0 < a < 1$. Updates estimate $\pi^t_i(k) = (1 - \lambda^t)\pi^t_i(k) + \lambda^t \pi$, where $\pi$ is profit for action $k$ in period $t$. Eventually converges, $\lambda^t = \beta \lambda^{t-1}$, $0 < \beta < 1$.

- Compare Groves, VCG and TG. (Plot mean absolute error $1/N \sum_i |k_i - k^*|$; efficiency $\sum_i v_i(x) / v(4)$.)

---

**Congestion Game**

- $N$ players. One bar: $y_i \in \{0, 1\}$ attend/not.

  $u_i(y) = xe^{y/\theta_i}$ if $y_i = 1$, and 0 otherwise, with $x = \sum y_i$. Type $\theta_i \in (1, 2, \ldots, 6)$ (higher, prefer less congestion.)

- Design goal: $\max_y \sum_i u_i(y)$ given $\theta = (\theta_1, \ldots, \theta_N)$.

- Compare Groves, VCG and TG. (Mean abs. error, and $\sum_i v_i(y) / \max_y \sum_i v_i(y)$.)

---

**On the Agenda in MAL**

(Shoham, Powers and Grenager)

Five possible agendas (and critique):

- **Descriptive**: How do Humans play?

- **Computational**: Learning to compute an equilibrium.

- **Prescriptive** (1): Program agents to succeed in converging to an optimal joint policy.

- **Prescriptive** (2): ("Equil. agenda") Agents self-interested. Bring the algorithms into an equilibrium.

- **Prescriptive** (3): ("AI agenda") Find the optimal single-agent algorithm, to perform against a fixed class of opponent algorithms (i.e. optimal given an environment.)

- **AI Agenda**: categorize agent strategies; empirical methods; $\epsilon$-optimal WRT target set & "security-level" against any opp.
Rationality and Bounded rationality

(Aumann 97)

- Economists are dissatisfied with strict rationality
  - human don’t fit model (introspection)
  - proven computational complexity
  - laboratory evidence
- "There is no unified theory of bounded rationality and probably never will be."
- Examines three different but related approaches.
- "to develop a meaningful definition of rationality in a situation in which calculation and analysis themselves are costly and/or limited."

Rationality and Bounded rationality

- Evolution as a BR model of game theory:
  - evolutionarily stable strategy (ESS); NE corresponds to population equilibria; predator and prey, bees and flowers, etc.)
  - evolutionary dynamics, myopic adjustment rules and study of convergent properties
  - rule rationality vs. act-rationality (divide money,... 50-50, 65-35, ...) “don’t be a sucker”
- Perturbations of rationality:
  - refinements, that note imperfections in play of others; or allow for “crazy” types
  - ϵ-equilibria: relaxed models of common knowledge
- Computability:
  - complexity of computing BR
  - games between automata, Turing machines
  - catalyst: Axelrod “prisoner’s dilemma” tournament, programs of a limited size in Fortran

Overview of Initial Lectures

[2 ] Game Theory I
- strategic form, Nash equil., dominant strategy, Bayesian-Nash.

[3 ] Game Theory II
- extensive form and repeated form games, folk theorems

- Revelation principle, Dominant strategy implementation

[5 ] Implementation Theory
- Nash implementation (complete information)

[6 ] Sequential decision making
- MDPs, Reinforcement learning

I: Multi-Agent Learning
(tentative... still moving around)


II: Implementation
(also tentative)

[16] COLlective INtelligence and Braess' Paradox by K.Tumer and D.Wolpert in JAIR 16:359-87, 2002


Next class

- Fast-Paced intro to Game Theory!