CAP 5993/CAP 4993
Game Theory

Instructor: Sam Ganzfried

sganzfri@cis.fiu.edu

Schedule

- HW4 due 4/13.
- Project presentations on 4/18 and 4/20.
- Project writeup due 4/20.
- Final exam on 4/25.

Game representations

- Strategic form
- Extensive form
  - Perfect information
  - Perfect information (with chance events)
  - Imperfect information (with chance events)
- Repeated
  - Finite vs. infinite
  - Discounted vs. undiscounted
- **Stochastic games**
- **Continuous games**
- **Bayesian games**

Solution concepts

- Maxmin strategies
- Weak/strict domination
- Nash equilibrium
- Refinements of Nash equilibrium
  - Trembling hand perfect equilibrium
  - Subgame perfect equilibrium
  - Proper equilibrium
  - Evolutionarily stable strategies
  - **Uniform equilibrium**
  - **Markov perfect equilibrium**
  - **Bayes-Nash equilibrium**
- Quantal response equilibrium
- Correlated equilibrium
- Stackelberg equilibrium

Stochastic (Markov) games

	L	R
T	0 s_2	1 s_2
B	1 s_1	0 s_0

State s_2

	L
T	1 s_1

State s_1

	L
T	0 s_0

State s_0

	L	R
T	$(1-\lambda)v_2$	$\lambda+(1-\lambda)v_2$
B	$\lambda+(1-\lambda)v_1$	$(1-\lambda)v_0$

Game $G_{s_2}^{\lambda}$

	L
T	$\lambda+(1-\lambda)v_1$

Game $G_{s_1}^{\lambda}$

	L
T	$(1-\lambda)v_0$

Game $G_{s_0}^{\lambda}$

- By imposing the fixed point condition on both states 0 and 1, we obtain: $v_0 = 0$, $v_1 = 1$.
- From the Indifference Principle, we can then solve for v_2 to obtain $v_2 = ½$.
- The equilibrium strategies are [½ L, ½ R] for row player and [$1/(1+ \lambda)$ T, $\lambda /(1+ \lambda)$ B] for column player.
- (Full derivation in Bauso textbook).

- Theorem (Shapley 1953): If all sets are finite, then for every λ there exists an equilibrium in stationary strategies.
  - Proof: Uses the above "dynamic programming" procedure, where "nonexpansiveness" of the value operator yields a unique fixed point, which corresponds to a Nash equilibrium.
- A strategy is *stationary* if it depends only on the current state (and not on the time step).

- Theorem (Mertens and Neyman 1981): For two-player zero-sum games, each player has a strategy that is ε-optimal for every discount factor sufficiently small.
  - Called a “uniform equilibrium”

- Theorem (Vielle 2000): For every two-player nonzero-sum stochastic game there is a strategy profile that is an ε-equilibrium for every discount factor sufficiently small.

- Open problem for more than two players.

Uniform equilibrium

- A strategy profile that is an ε-equilibrium for every T sufficiently large and for every λ sufficiently small is called a **uniform ε-equilibrium**.

- Definition: Let $\varepsilon > 0$. A strategy profile σ is a **uniform ε-equilibrium** if there are T_0 in N and λ_0 in (0,1) such that for every $T >= T_0$ the strategy profile σ is a T-stage ε-equilibrium, and for every λ in $(0, \lambda_0)$ it is a λ-discounted ε-equilibrium.

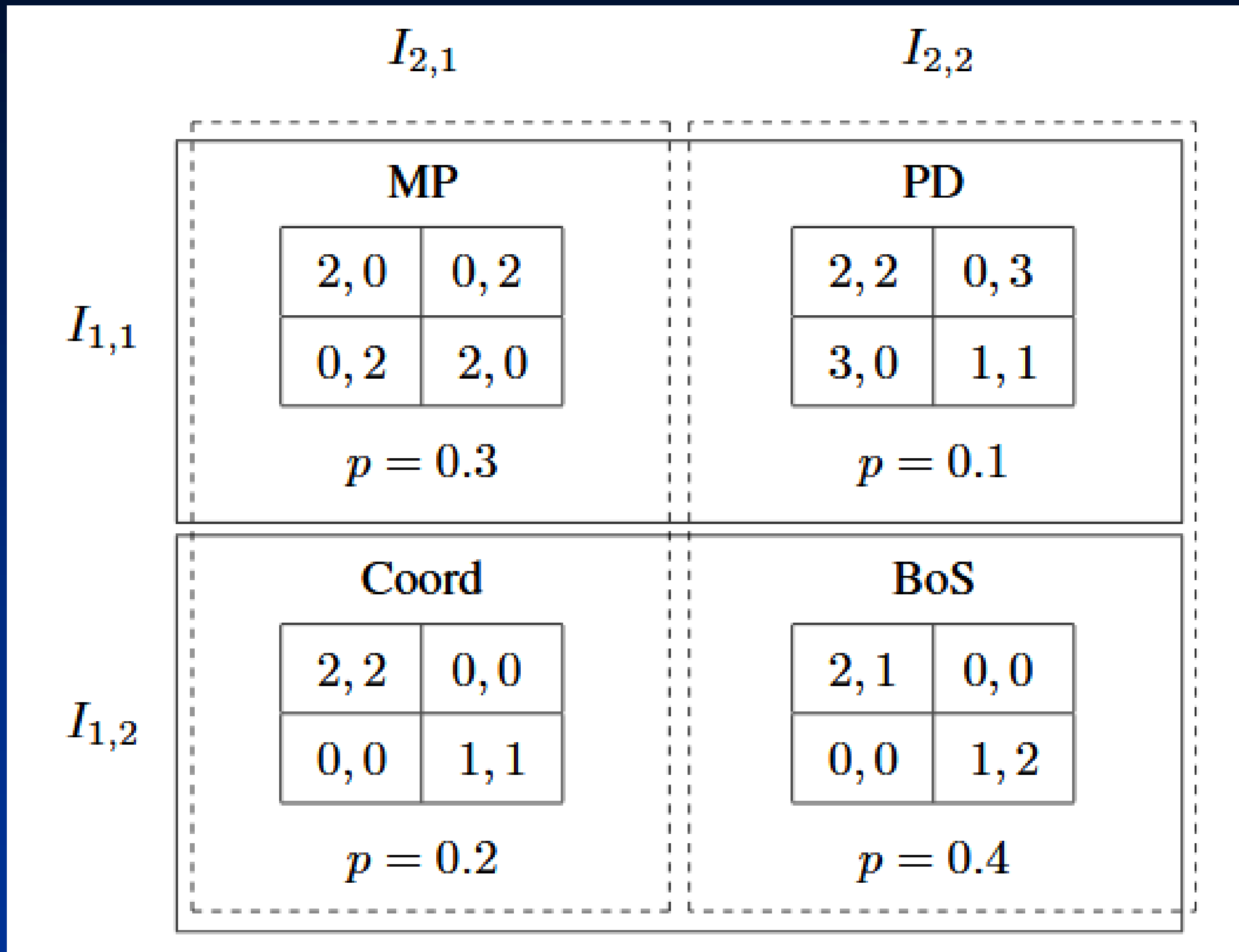
- If for every $\varepsilon > 0$ the game has a uniform ε-equilibrium with corresponding payoff gε, then any accumulation point of $(g_\varepsilon)_{\varepsilon>0}$ as ε goes to 0 is a **uniform equilibrium payoff**.

- http://www.math.tau.ac.il/~eilons/encyclopedia.pdf

Markov perfect equilibrium

- In extensive form games, and specifically in stochastic games, a Markov perfect equilibrium is a set of mixed strategies for each of the players which satisfy the following criteria:
  - The strategies have the Markov property of memorylessness, meaning that each player's mixed strategy can be conditioned only on the *state* of the game. These strategies are called *Markov reaction functions*.
  - The *state* can only encode payoff-relevant information. This rules out strategies that depend on non-substantive moves by the opponent. It excludes strategies that depend on signals, negotiation, or cooperation between the players (e.g. cheap talk or contracts).
  - The strategies form a subgame perfect equilibrium of the game.
- Markov perfect equilibrium is a refinement of the concept of subgame perfect equilibrium to extensive form games

Bayesian games



a_1	a_2	θ_1	θ_2	u_1	u_2
U	L	$\theta_{1,1}$	$\theta_{2,1}$	2	0
U	L	$\theta_{1,1}$	$\theta_{2,2}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,1}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,2}$	2	1
U	R	$\theta_{1,1}$	$\theta_{2,1}$	0	2
U	R	$\theta_{1,1}$	$\theta_{2,2}$	0	3
U	R	$\theta_{1,2}$	$\theta_{2,1}$	0	0
U	R	$\theta_{1,2}$	$\theta_{2,2}$	0	0

a_1	a_2	θ_1	θ_2	u_1	u_2
D	L	$\theta_{1,1}$	$\theta_{2,1}$	0	2
D	L	$\theta_{1,1}$	$\theta_{2,2}$	3	0
D	L	$\theta_{1,2}$	$\theta_{2,1}$	0	0
D	L	$\theta_{1,2}$	$\theta_{2,2}$	0	0
D	R	$\theta_{1,1}$	$\theta_{2,1}$	2	0
D	R	$\theta_{1,1}$	$\theta_{2,2}$	1	1
D	R	$\theta_{1,2}$	$\theta_{2,1}$	1	1
D	R	$\theta_{1,2}$	$\theta_{2,2}$	1	2

Figure 6.9 Utility functions u_1 and u_2 for the Bayesian game from Figure 6.7.

Our third definition uses the notion of an *epistemic type*, or simply a *type* as a way of defining uncertainty directly over a game's utility function.

Definition 6.3.2 (Bayesian game: types) *A* Bayesian game *is a tuple* (N, A, Θ, p, u) *where:*

- *N is a set of agents;*
- *$A = A_1 \times \cdots \times A_n$, where A_i is the set of actions available to player i;*
- *$\Theta = \Theta_1 \times \ldots \times \Theta_n$, where Θ_i is the type space of player i;*
- *$p : \Theta \mapsto [0, 1]$ is a common prior over types; and*
- *$u = (u_1, \ldots, u_n)$, where $u_i : A \times \Theta \mapsto \mathbb{R}$ is the utility function for player i.*

The assumption is that all of the above is common knowledge among the players, and that each agent knows his own type. This definition can seem mysterious, because the notion of type can be rather opaque. In general, the type of agent encapsulates all the information possessed by the agent that is not common knowledge. This is often quite simple (e.g., the agent's knowledge of his private payoff function), but can also include his beliefs about other agents' payoffs, about their beliefs about his own payoff, and any other higher-order beliefs.

We can get further insight into the notion of a type by relating it to the formulation at the beginning of this section. Consider again the Bayesian game in Figure 6.7. For each of the agents we have two types, corresponding to his two information sets. Denote player 1's actions as U and D, player 2's actions as L and R. Call the types of the first agent $\theta_{1,1}$ and $\theta_{1,2}$, and those of the second agent $\theta_{2,1}$ and $\theta_{2,2}$. The joint distribution on these types is as follows: $p(\theta_{1,1}, \theta_{2,1}) = .3$, $p(\theta_{1,1}, \theta_{2,2}) = .1$, $p(\theta_{1,2}, \theta_{2,1}) = .2$, $p(\theta_{1,2}, \theta_{2,2}) = .4$. The conditional probabilities for the first player are $p(\theta_{2,1} \mid \theta_{1,1}) = 3/4$, $p(\theta_{2,2} \mid \theta_{1,1}) = 1/4$, $p(\theta_{2,1} \mid \theta_{1,2}) = 1/3$, and $p(\theta_{2,2} \mid \theta_{1,2}) = 2/3$. Both players' utility functions are given in Figure 6.9.

- Recall that in an imperfect-information extensive-form game a pure strategy is a mapping from information sets to actions. The definition is similar in Bayesian games: a pure strategy is a mapping from every *type* the agent could have to the action he would play if he had that type.

Bayes-Nash equilibrium

- In a non-Bayesian game, a strategy profile is a Nash equilibrium if every strategy in that profile is a best response to every other strategy in the profile; i.e., there is no strategy that a player could play that would yield a higher payoff, given all the strategies played by the other players. In a Bayesian game (where players are modeled as risk-neutral), rational players are seeking to maximize their expected payoff, given their beliefs about the other players (in the general case, where players may be risk-averse or risk-loving, the assumption is that players are expected utility-maximizing).
- A *Bayesian Nash equilibrium* is defined as a strategy profile and *beliefs specified for each player about the types of the other players* that maximizes the expected payoff for each player given their beliefs about the other players' types and given the strategies played by the other players.

Robust management of diabetes

- A finite horizon Markov decision process is a tuple M = (S, A, R, P H).
  - S is finite set of states
  - A is finite set of actions
  - H is the horizon
  - System starts in initial state s_0
  - P(s,a,s’) is probability of transitioning from state s to state s’ after taking action a from A.
  - The immediate reward from taking action a at state s is R(s,a).
- Objective is to compute policy π that maximizes the expected cumulative reward.

- Markov randomized policy: those that map only the current state and timestep to a probability distribution over actions.
- For a fixed MDP M, the set of Markov random policies (in fact, Markov deterministic policies), contains a maximizing policy. This is called the optimal policy for the fixed MDP. However, for MDPs with parameter uncertainty, Markov random policies may not be a sufficient class.
- MDPs with parameter uncertainty: R and P are not known.

Diabetes management

- MDP with $|S| = 9$ states, $|A| = 3$ actions, and a time horizon of $H = 3$. States are a combination of blood glucose level and meal size. Three times daily, corresponding to meal times, the patient injects themselves with a dose of insulin to bring down the rise in blood glucose in the moderate range all day. The uncertain reward function is sampled from a independent multivariate Normal distribution and transition probabilities are sampled from Dirichlet distributions.

CFR-BR

- Can solve for optimal robust policies using "CFR-BR."
- CFR: counterfactual regret minimization algorithm (plays two regret minimization algorithms against each other and converges to Nash equilibrium in the limit).
- Variant CFR-BR addresses the challenge of an adversary having an intractably large strategy space. Combines two ideas:
  - Avoids representing entirety of second player's strategy space by having the player always play according to a best response to the first player's strategy.
  - Avoids having to compute or store a complete best-response by sampling over chance outcomes to focus the best-response and regret updates on small subtree of the game on each iteration. This removes dependence on the size of the adversary's strategy space in either computation time or memory. Also has convergence guarantee to Nash equilibrium.

Open-bid ascending auction (English auction)

- This is the most common public auction. It is characterized by an auctioneer who publicly declares the price of the object offered for sale. The opening price is low, and as long as there are at least two buyers willing to pay the declared price, the auctioneer raises the price (either in discrete jumps, or in a continuous manner using a clock). Each buyer raises a hand as long as he is willing to pay the last price that the auctioneer has declared. The auction ends when all hands except one have been lowered, and the object is sold to the last buyer whose hand is still raised, at the last price declared by the auctioneer. If the auction ends in a draw, a previously agreed rule (such as tossing a coin) is employed to determine who wins the object, which is then sold to the winner at the price that was current when they lowered their hands.
  - Used for web-based auctions and auctions of works of art.
  - https://www.youtube.com/watch?v=UJafks_lAc8

Open-bid descending auction (Dutch auction)

- The auctioneer begins by declaring a very high price, higher than any buyer could be expected to pay. As long as no buyer is willing to pay the last declared price, the auctioneer lowers the declared price (either in discrete jumps or in a continuous manner using a clock), up to the point at which at least one buyer is willing to pay the declared price and indicates his readiness by raising his hand or pressing a button to stop the clock. If the price drops below a previously declared minimum, the auction is stopped, and the object on offer is not sold. Similarly to the English auction, a previously agreed rule is employed to determine who wins the auction if two or more buyers stop the clock at the same time.
  - The flower auction at the Aalsmeer Flower Exchange, near Amsterdam, is conducted using this method.
  - https://www.youtube.com/watch?v=SkehTvMHiiw

Sealed-bid first-price auction

- In this method, every buyer in the auction submits a sealed envelope containing the price he is willing to pay for the offered object. After all buyers have submitted their offers, the auctioneer opens the envelopes and reads the offers they contain. The buyer who has submitted the highest bid wins the offered object, and pays the price that he has bid. A previously agreed rule determines how to resolve draws.

- Theorem: The sealed-bid first price auction is equivalent to the open-bid descending auction.

Sealed-bid second price auction (Vickrey auction)

- The sealed-bid second-price auction method is similar to the first-price sealed-bid auction method, except that the winner of the auction, i.e., buyer who submitted the highest bid, pays the *second*-highest price among the bid prices for the offered object. A previously agreed-upon rule determines the winner in case of a draw, with a winner in this case paying what he bid (which is, in the case of a draw, also the second-highest bid).

Assignment

- Reading for next class: chapter 21 from main textbook.
- Homework 4 out (due 4/13).