Last update: February 10, 2009

Non-Zero-Sum Games

Dana Nau, CMSC 828N

Introduction

We'll now look at games that may have
- imperfect information
- nonzero-sum outcomes

Remember we defined a strategy as a *policy*, i.e., a function from states into actions that tells what you will do in every game state where it's your move.

The Prisoner's Dilemma

Scenario: The police arrest two suspects for a crime. They tell each one they'll reduce his/her prison sentence if he/she testifies against the other.

Each prisoner must choose between two actions:

- cooperate with the other prisoner, i.e., don't betray him/her
- defect (betray the other prisoner).

Payoff $= -$(Years in prison):

		P_2 C	P_2 D
P_1	C	$-2, -2$	$-5, 0$
P_1	D	$0, -5$	$-4, -4$

Nonnegative payoffs:

		P_2 C	P_2 D
P_1	C	$3, 3$	$0, 5$
P_1	D	$5, 0$	$1, 1$

◇ Each strategy is a single action (cooperate or defect)

◇ Payoffs are non-zero-sum (i.e., non-constant-sum)

◇ Imperfect information: neither player knows the other's move until after *both* players have moved

Strategies

Suppose there are players $P_1, \ldots, P_n$
For each i, let $S_i = \{$all possible strategies for $P_i\}$
Notation: s_i refers to a strategy in S_i

Strategy profile: an n-tuple $(s_1, s_2, \ldots, s_n)$, one strategy for each player

Utility $U_{P_i}(s_1, \ldots, s_n) = P_i$'s expected payoff given $(s_1, s_2, \ldots, s_n)$

s_i *strongly dominates* s_i' if P_i always does better with s_i than with s_i':

$$\forall\ s_1, \ldots, s_{i-1},\ s_{i+1}, \ldots, s_n,$$
$$U_{P_i}(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n) > U_{P_i}(s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_n)$$

s_i *weakly dominates* s_i' if P_i never does worse with s_i than with s_i', and there is at least one case where P_i does better with s_i than with s_i':

$$\forall\ s_1, \ldots, s_{i-1},\ s_{i+1}, \ldots, s_n \quad U_{P_i}(\ldots, s_i, \ldots) \geq U_{P_i}(\ldots, s_i', \ldots)$$

and

$$\exists\ s_1, \ldots, s_{i-1},\ s_{i+1}, \ldots, s_n \quad U_{P_i}(\ldots, s_i, \ldots) > U_{P_i}(\ldots, s_i', \ldots)$$

Dominant strategy equilibrium

s_i is a *(strongly, weakly) dominant* strategy if it (strongly, weakly) dominates every $s'_i \in S_i$.

Dominant strategy equilibrium: a strategy profile $(s_1, \ldots, s_n)$ such that $\forall i$, s_i is strongly dominant for P_i

i.e., for every possible set of strategies the other players might use, P_i does best by using s_i

Example: the Prisoner's Dilemma

		P_2 C	P_2 D
P_1	C	$3, 3$	$0, 5$
	D	$5, 0$	$1, 1$

Dominant strategy equilibrium: (D,D)

Can there ever be more than one dominant strategy equilibrium?

Pareto optimality

Let $S = (s_1, s_2, \ldots, s_n)$ and $S' = (s'_1, s'_2, \ldots, s'_n)$ be strategy profiles.

S *Pareto dominates* S' if $\forall i, U_{P_i}(S) \geq U_{P_i}(S')$ and $\exists i \; U_{P_i}(S) > U_{P_i}(S')$

S is *Pareto optimal* if no strategy profile Pareto dominates S

E.g., consider the Prisoner's Dilemma again:

		P_2 C	P_2 D
P_1	C	$3, 3$	$0, 5$
P_1	D	$5, 0$	$1, 1$

- ♢ The dominant strategy equilibrium, (D,D), is **not** Pareto optimal: (C,C) Pareto dominates it
- ♢ (D,C) and (C,D) are both Pareto optimal
- ♢ Is (C,C) Pareto optimal?

Dominant strategy equilibrium (continued)

Not every game has a dominant strategy equilibrium

Example:

Acme and *Best* are two video-game manufacturers
Each needs to decide whether to use DVDs or CDs in its next game machine

		Best	
		DVD	CD
Acme	DVD	$9,\ \ 9$	$-4,-1$
	CD	$-3,-1$	$5,\ \ 5$

Nash equilibrium

A *Nash equilibrium* is basically a local optimum:
a strategy profile $(s_1, \ldots, s_n)$ such that no player can benefit from switching to a different strategy if nobody else switches

I.e., $\forall i, \ \forall \ s'_i \in S_i$

$$U_{P_i}(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n) \geq U_{P_i}(s_1, \ldots, s_{i-1}, s'_i, s_{i+1}, \ldots, s_n)$$

Every dominant strategy equilibrium is a Nash equilibrium, but not vice versa
Every game has at least one Nash equilibrium

		Best	
		DVD	CD
Acme	DVD	$9, \ \ 9$	$-4, -1$
	CD	$-3, -1$	$5, \ \ 5$

Two Nash equilibria: (DVD, DVD) and (CD, CD)
One of them is Pareto optimal

Mixed strategies

Two-finger Morra:
Two players: Odd and Even.
Each holds up 1 or 2 fingers:

		Odd 1	Odd 2
Even	1	$2, -2$	$-3, \ 3$
Even	2	$-3, \ 3$	$4, -4$

Among *pure* (deterministic) strategies, there is no Nash equilibrium

- If Odd chooses 1, Even can win by choosing 1
- If Even chooses 1, Odd can win by choosing 2

The Nash equilibrium uses *mixed* (randomized) strategies.

What are they?

Mixed strategies

Two-finger Morra:

Two players: Odd and Even.

Each holds up 1 or 2 fingers:

		Odd 1	Odd 2
Even	1	$2, -2$	$-3, 3$
Even	2	$-3, 3$	$4, -4$

This is a zero-sum game, so the Minimax theorem applies:

The game has a minimax value V.

Odd has a strategy t that guarantees $U_{\text{Even}} \leq V$.

Even has a strategy s that guarantees $U_{\text{Even}} \geq V$.

s and t are a Nash equilibrium.

Can't use the minimax algorithm to find s and t:
It's only for finding pure strategies on perfect-information games

Instead, use von Neumann's **maximin technique** ...

Modify the game to give Odd an advantage

Make Even choose a strategy first, and tell Odd what it is.
Even's strategy is $[P(1) = p, P(2) = 1 - p]$ for some p.

◊ If Odd's move is 1, Even's expected utility is $2p - 3(1 - p) = 5p - 3$.
◊ If Odd's move is 2, Even's expected utility is $-3p + 4(1 - p) = 4 - 7p$.
◊ If Odd's strategy is mixed, Even's expected utility is in between

Odd's best strategy: choose the action that minimizes Even's expected utility. In this case, Even's expected utility is:

$$U_{\text{Even}|p} = \min(5p - 3, 4 - 7p)$$

so Even's best strategy is to choose p that maximizes $U_{\text{Even}|p}$
This occurs where the line $y = 5p - 3$ intersects the line $y = 4 - 7p$

$$5p - 3 = 4 - 7p \quad \Rightarrow \quad 12p = 7 \quad \Rightarrow \quad p = 7/12$$

Expected utility for Even in this case is $-1/12$
This is a lower bound on Even's best expected utility in Two-finger Morra

Try giving Even the advantage instead

Make Odd choose a strategy first, and tell Even what it is.
Odd's strategy is $[P(1) = q, P(2) = 1 - q]$ for some q.

- ♢ If Even's move is 1, Even's expected utility is $2q - 3(1 - q) = 5q - 3$.
- ♢ If Even's move is 2, Even's expected utility is $-3q + 4(1 - q) = 4 - 7q$.
- ♢ If Even's strategy is mixed, Odd's expected utility is in between

Even's best strategy: choose the action that maximizes Even's expected utility. In this case, Even's expected utility is

$$U_{\text{Even}|q} = \max(5q - 3, 4 - 7q)$$

so Odd's best strategy is to choose q that minimizes $U_{\text{Even}|q}$
This occurs where the line $y = 5q - 3$ intersects the line $y = 4 - 7q$

$$5q - 3 = 4 - 7q \quad \Rightarrow \quad 12q = 7 \quad \Rightarrow \quad q = 7/12$$

Expected utility for Even in this case is $-1/12$
This is an upper bound on Even's best expected utility in Two-finger Morra

Back to the unmodified game

Two-finger Morra:
Two players: Odd and Even.
Each holds up 1 or 2 fingers:

		Odd	
		1	2
Even	1	$2, -2$	$-3, \ \ 3$
	2	$-3, \ \ 3$	$4, -4$

We have determined that

Lower bound on Even's best expected utility is $-1/12$
occurs when $p = 7/12$

Upper bound on Even's expected utility is also $-1/12$
occurs when $q = 7/12$

Thus the game's minimax value is $-1/12$

The minimax value occurs when Even chooses $p = 7/12$ and Odd chooses $q = 7/12$

Nash equilibria in other cases

How to find Nash equilibria in other cases?

There are techniques, but I haven't studied them

Does anyone want to do a classroom presentation on this?

Prisoner's Dilemma

		P_2	
		C	D
P_1	C	$3, 3$	$0, 5$
	D	$5, 0$	$1, 1$

(D,D) is a dominant equilibrium, but it isn't Pareto optimal

(C,C) is Pareto optimal, but it's not even a Nash equilibrium

How to get both players to choose C?

- ♢ Each player must be willing to forego the personal gain that he/she would get from defecting
- ♢ Each player has to trust the other to do the same

How to make this happen?

Iterated Prisoner's Dilemma

		P_2	
		C	D
P_1	C	$3, 3$	$0, 5$
	D	$5, 0$	$1, 1$

Iterated Prisoner's Dilemma: play the Prisoner's Dilemma repeatedly, score is the sum of the payoffs in all the iterations

- ◇ If you defect and they cooperate, you get a short-term gain
  But they might punish you next time by defecting
- ◇ You can both do well if you both cooperate with each other
- ◇ How to establish and maintain cooperation,
  without letting them take advantage of you?

Some well-known strategies

- Tit for Tat (TFT):
  Move 1: cooperate
  Move i: do what the other
  player did on move $i-1$

Iteration	TFT	other player
1	C	C
2	C	D
3	D	C
4	C	C

- Tit for Two Tats: cooperate unless the other player defected twice
- GRIM: if the other player ever defects, never cooperate again
- AllC: always cooperate
- AllD (the hawk strategy): always defect
- Tester: Defect on move 1. If the opponent retaliates, then play Tit-for-Tat. Otherwise intersperse cooperation and defection

TFT with other players

Axelrod's famous tournaments

- ♢ [Axelrod, The Evolution of Cooperation, 1985]
- ♢ In these tournaments, TFT usually did best
- ♢ It could establish and maintain cooperations with many other players
- ♢ It could prevent malicious players from taking advantage of it

TFT	AllC	TFT	AllD	TFT	Grim	TFT	TFT	TFT	Tester
C	C	C	D	C	C	C	C	C	D
C	C	D	D	C	C	C	C	D	C
C	C	D	D	C	C	C	C	C	C
C	C	D	D	C	C	C	C	C	C
C	C	D	D	C	C	C	C	C	C
C	C	D	D	C	C	C	C	C	C
C	C	D	D	C	C	C	C	C	C
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Axelrod also looked at analogies with various human behaviors

Example: trench warfare in World War I

Incentive to cooperate:

◇ If I attack the other side, then they'll retaliate and I'll get hurt

◇ If I don't attack, maybe they won't either

Result: tacit cooperation

◇ Even though the soldiers were supposed to be enemies, they tried to avoid attacking each other

◇ This was one reason (though certainly not the only one) why World War I lasted so long