Game Theory: Normal Form Games

CPSC 322 Lecture 34

April 3, 2006

Reading: excerpt from "Multiagent Systems", chapter 3.

Lecture Overview

Rewards and Values

Suppose the agent receives the sequence of rewards $r_1, r_2, r_3, r_4, \ldots$. What value should be assigned?

- total reward $V = \sum_{i=1}^{\infty} r_i$
- average reward $V = \lim_{n \to \infty} \frac{r_1 + \cdots + r_n}{n}$
- discounted reward $V = \sum_{i=1}^{\infty} \gamma^{i-1} r_i$
  - γ is the discount factor
  - $0 \leq \gamma \leq 1$

Policies

- A stationary policy is a function:

$$\pi : S \to A$$

  Given a state s, $\pi(s)$ specifies what action the agent who is following π will do.
- An optimal policy is one with maximum expected value
  - we'll focus on the case where value is defined as discounted reward.
- For an MDP with stationary dynamics and rewards with infinite or indefinite horizon, there is always an optimal stationary policy in this case.

Value of a Policy

- $Q^\pi(s,a)$, where a is an action and s is a state, is the expected value of doing a in state s, then following policy π.
- $V^\pi(s)$, where s is a state, is the expected value of following policy π in state s.
- Q^π and V^π can be defined mutually recursively:

$$\begin{aligned} V^\pi(s) &= Q^\pi(s,\pi(s)) \\ Q^\pi(s,a) &= \sum_{s'} P(s'|a,s)\left(r(s,a,s') + \gamma V^\pi(s')\right) \end{aligned}$$

Value of the Optimal Policy

- $Q^*(s,a)$, where a is an action and s is a state, is the expected value of doing a in state s, then following the optimal policy.
- $V^*(s)$, where s is a state, is the expected value of following the optimal policy in state s.
- Q^* and V^* can be defined mutually recursively:

$$\begin{aligned} Q^*(s,a) &= \sum_{s'} P(s'|a,s)\left(r(s,a,s') + \gamma V^*(s')\right) \\ V^*(s) &= \max_a Q^*(s,a) \\ \pi^*(s) &= \arg\max_a Q^*(s,a) \end{aligned}$$

Value Iteration

- Idea: Given an estimate of the k-step lookahead value function, determine the $k+1$ step lookahead value function.
- Set V_0 arbitrarily.
  - e.g., zeros
- Compute Q_{i+1} and V_{i+1} from V_i:

$$
\begin{aligned}
Q_{i+1}(s,a) &= \sum_{s'} P(s'|a,s)\left(r(s,a,s') + \gamma V_i(s')\right) \\
V_{i+1}(s) &= \max_a Q_{i+1}(s,a)
\end{aligned}
$$

- If we intersect these equations at Q_{i+1}, we get an update equation for V:

$$V_{i+1}(s) = \max_a \sum_{s'} P(s'|a,s)\left(r(s,a,s') + \gamma V_i(s')\right)$$

Asynchronous VI: storing $Q[s, a]$

- Repeat forever:
  - Select state s, action a;
  - $Q[s,a] \leftarrow \sum_{s'} P(s'|s,a) \left(R(s,a,s') + \gamma \max_{a'} Q[s',a'] \right);$

Lecture Overview

Recap

Game Theory

Example Matrix Games

Non-Cooperative Game Theory

- What is it?

Non-Cooperative Game Theory

- What is it?
  - mathematical study of interaction between rational, self-interested agents

Non-Cooperative Game Theory

- What is it?
  - mathematical study of interaction between rational, self-interested agents
- Why is it called non-cooperative?

Non-Cooperative Game Theory

- What is it?
  - mathematical study of interaction between rational, self-interested agents
- Why is it called non-cooperative?
  - while it's most interested in situations where agents' interests conflict, it's not restricted to these settings
  - the key is that the individual is the basic modeling unit, and that individuals pursue their own interests
    - cooperative/coalitional game theory has teams as the central unit, rather than agents
- You can think of a non-cooperative game as a decision diagram where different agents control different decision nodes, and where each agent has his own utility node.

TCP Backoff Game

Should you send your packets using correctly-implemented TCP (which has a "backoff" mechanism) or using a defective implementation (which doesn't)?

- Consider this situation as a two-player game:
  - both use a correct implementation: both get 1 ms delay
  - one correct, one defective: 4 ms delay for correct, 0 ms for defective
  - both defective: both get a 3 ms delay.

TCP Backoff Game

- Consider this situation as a two-player game:
  - both use a correct implementation: both get 1 ms delay
  - one correct, one defective: 4 ms delay for correct, 0 ms for defective
  - both defective: both get a 3 ms delay.
- Questions:
  - What action should a player of the game take?
  - Would all users behave the same in this scenario?
  - What global patterns of behaviour should the system designer expect?
  - Under what changes to the delay numbers would behavior be the same?
  - What effect would communication have?
  - Repetitions? (finite? infinite?)
  - Does it matter if I believe that my opponent is rational?

Defining Games

- Finite, n-person game: $\langle N, A, u \rangle$:
  - N is a finite set of n players, indexed by i
  - $A = A_1, \ldots, A_n$ is a set of actions for each player i
    - $a \in A$ is an action profile
  - $u = \{u_1, \ldots, u_n\}$, a utility function for each player, where $u_i : A \mapsto \mathbb{R}$

- Writing a 2-player game as a matrix:
  - row player is player 1, column player is player 2
  - rows are actions $a \in A_1$, columns are $a' \in A_2$
  - cells are outcomes, written as a tuple of utility values for each player

Lecture Overview

Recap

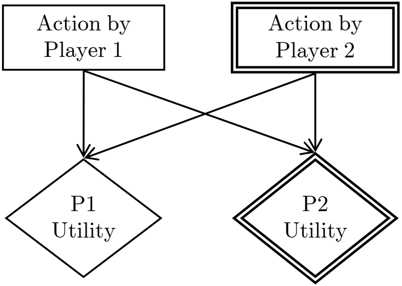
Game Theory

Example Matrix Games

Games in Matrix Form

Here's the TCP Backoff Game written as a matrix ("normal form") and as a decision network.

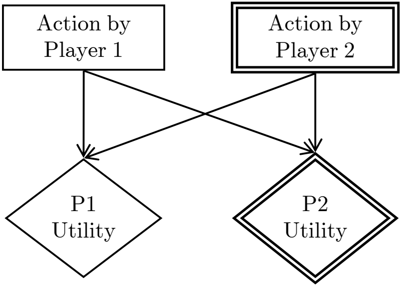
	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$



Games in Matrix Form

Here's the TCP Backoff Game written as a matrix ("normal form") and as a decision network.

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$



Play this game with someone near you, repeating five times.

More General Form

Prisoner's dilemma is any game

	C	D
C	a, a	b, c
D	c, b	d, d

with $c > a > d > b$.

Games of Pure Competition

Players have exactly opposed interests

- There must be precisely two players (otherwise they can't have exactly opposed interests)
- For all action profiles $a \in A$, $u_1(a) + u_2(a) = c$ for some constant c
  - Special case: zero sum
- Thus, we only need to store a utility function for one player

Matching Pennies

One player wants to match; the other wants to mismatch.

	Heads	Tails
Heads	1	-1
Tails	-1	1

Matching Pennies

One player wants to match; the other wants to mismatch.

	Heads	Tails
Heads	1	-1
Tails	-1	1

Play this game with someone near you, repeating five times.

Rock-Paper-Scissors

Generalized matching pennies.

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

...Believe it or not, there's an annual international competition for this game!

Games of Cooperation

Players have exactly the same interests.

- no conflict: all players want the same things
- $\forall a \in A, \forall i, j,\ u_i(a) = u_j(a)$
- we often write such games with a single payoff per cell
- why are such games "noncooperative"?

Coordination Game

Which side of the road should you drive on?

	Left	Right
Left	1	0
Right	0	1

Coordination Game

Which side of the road should you drive on?

	Left	Right
Left	1	0
Right	0	1

Play this game with someone near you, repeating five times.

General Games: Battle of the Sexes

The most interesting games combine elements of cooperation *and* competition.

	B	F
B	2, 1	0, 0
F	0, 0	1, 2

General Games: Battle of the Sexes

The most interesting games combine elements of cooperation *and* competition.

	B	F
B	$2, 1$	$0, 0$
F	$0, 0$	$1, 2$

Play this game with someone near you, repeating five times.