Multi-agent AI: Game Theory

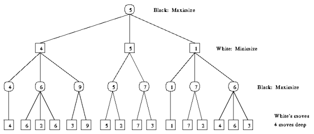
Sections 17.6-17.7

What is different about multi-agent environments?

- Single-agent: decision-making and acting "against" or "with" nature
  - indifferent to agent's goals
- Multi-agent: against or with (an)other agent(s)
  - have a stake in agent's goals
  - possibly require coordination of plans/policies

Turn-taking games

- Minimax



- Expectiminimax

Game theory

- Simultaneous moves allowed
  - equivalently, agent A does not know what agent B's last move is

Game types

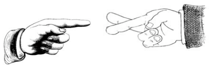
- Cooperative vs. competitive (zero-sum or general-sum)
- Perfect vs. imperfect information
- One-move vs. sequential
- Discrete vs. continuous
- Symmetric vs. asymmetric (does utility of a state depend on who is playing?)

Today

- Multi-agent decision-making (under uncertainty)
  - game theory
- Designing the rules of a game to achieve some overarching goal
  - mechanism design
- Sequential cooperative games

Game theory basics

- Game: two-finger Morra



- Players
- Actions
- Payoff matrix

	O: one	O: two
E: one	E=2, O=-2	E=-3, O=3
E: two	E=-3, O=3	E=4, O=-4

- Strategy (= policy)
  - **pure strategy** (= deterministic policy)
  - **mixed strategy** (= stochastic policy)
- **Outcome** (= result of a strategy profile) vs. **solution** (= rational strategies for each player)

Prisoner's Dilemma

"Two suspects are arrested by the police. The police have insufficient evidence for a conviction, and, having separated both prisoners, visit each of them to offer the same deal. If one testifies ("defects") for the prosecution against the other and the other remains silent ("cooperates"), the betrayer goes free and the silent accomplice receives the full 10-year sentence. If both remain silent, both prisoners are sentenced to only one year in jail for a minor charge. If each betrays the other, each receives a five-year sentence. Each prisoner must choose to betray the other or to remain silent. Each one is assured that the other would not know about the betrayal before the end of the investigation. How should the prisoners act?"

	Alice: defect	Alice: cooperate
Bob: defect	-5, -5	0, -10
Bob: cooperate	-10, 0	-1, -1

Dominance

- What should (rational) Alice do?
  - suppose Bob defects, then Alice should defect too
  - suppose Bob cooperates, then Alice should defect
- Rational strategy = always defect (**dominant strategy**)
- Dilemma: dominant strategy leads to a suboptimal outcome

	Alice: defect	Alice: cooperate
Bob: defect	-5, -5	0, -10
Bob: cooperate	-10, 0	-1, -1

Nash equilibrium


John Nash, b. 1928

- Strategies are in a Nash equilibrium if no player can gain by unilaterally changing their strategy

game solution

- What are the equilibrium points, if any, in Prisoner's Dilemma?
- In two-finger Morra?

	Alice: defect	Alice: cooperate
Bob: defect	-5, -5	0, -10
Bob: cooperate	-10, 0	-1, -1

No pure strategy equilibrium

- Consider mixed (stochastic) strategies
- E plays mixed strategy [*p:one, (1-p):two*]
- Expected utility = *p u(one)* + *(1-p) u(two)*

- *U(E,O)* and *U(O,E)*
- Mixed strategy equilibrium point for $p=7/12$ is the **maximum equilibrium**
  - (J. von Neumann's general result for every two-player zero-sum game)
- So, how should the rational agent play?

Repeated games

- Iterated Prisoner's Dilemma:
  - Alice and Bob will meet again (and play the game again) for some **unknown** number of times

- Does this change the dominant strategy?

- What if the number of repeated PD games is known?

Applications of PD

Should I take a personal cost for the greater good of the community?

- Kant's categorical imperative
- The tragedy of the commons. G. Hardin, *Science* 1968
- Climate change
- Arms race
- Sports: drafting
- Economics: advertising and cartels

	defect	**cooperate**
defect	lose, lose	win much, lose much
cooperate	lose much, win much	win, win

What about AI?

- Automating decision-making in game situations (e.g., in economics)
- CS-specific: traffic collision problems (e.g., packet collision in Ethernet networks)

Mechanism design

- Metagame: maximize the utility of the game rules
- Given that agents are rational, what game should we design…?
  - …s.t. its solution, with individual agents pursuing rational self-interested strategies, maximizes some **global utility function**
- Inverse game theory

Examples

- Routing TCP packets
- Distributed scheduling, task selection or load allocation
- Assigning medical interns to hospitals
- Auctioning off cheap airline tickets

Mechanism basics

- A mechanism consists of
  1. A language describing the set of **allowable** strategies
  2. An **outcome rule** G = payoff to agents given strategy profile

- Mechanism assigns an explicit cost (price) to all **externalities** (effects on global utility not recognized by individual utilities)
  - set costs (prices) s.t. rational agents acting in self-interest will maximize global utility (aka **social welfare**)

Auctions

- Private information
  - each agent has some private value v_i for the goods
- Each agent can place a bid b_i
- Highest bid takes the goods at some price (need not be equal to the bid)

Auction types

- English auction
  - auctioneer increments price of goods by some amount *d* until only one bidder is left; winner pays b_m+d
  - dominant bidding strategy is to stay in the game as long as current price is below your private value
  - **strategy-proof mechanism**: dominant strategy is to reveal your actual private value
  - high communication cost
- Sealed bid auction
  - each agent makes a single bid; highest bid wins
  - revealing true value is no longer dominant
  - best strategy depends on what you think others will bid

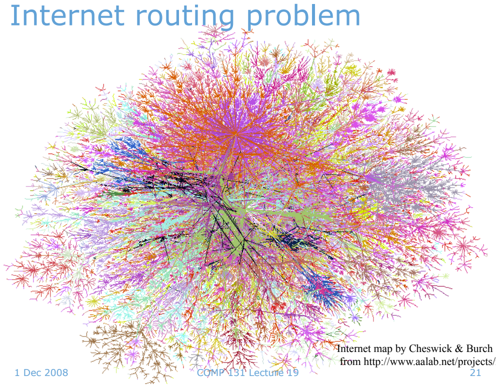
Vickrey auction

- Sealed bid second-price auction
  - highest bidder pays the second-highest price
- Game between you and highest other bidder

Utility of bid for agent i $u_i(b_i,b_m) = \begin{cases} (v_i - b_m) & \text{if } b_i > b_m \\ 0 & \text{otherwise} \end{cases}$

  - v_i is the only optimal bid for all values of b_m
  - revealing actual private value is the dominant strategy

Internet routing problem



Internet map by Cheswick & Burch
from http://www.aalab.net/projects/maps/

Internet routing problem

- Players are edges in network graph
- Individual edge costs c_i
- Minimize total path cost
  - players need to report their true cost, but may wish to report higher cost to avoid traffic
- Need strategy-proof mechanism
  - pay each player a payoff

$$p_i = length(path\ with\ c_i = \infty) - length(path\ with\ c_i = 0)$$

  - can prove that dominant strategy = reveal true cost
  - high communication and central computation cost

Cooperative games

- A class of problems in multi-agent AI can be described as cooperative games
  - cooperation always induces greater utilities, defection always decreases the utility
- Decentralized control of robot teams
- Network routing optimization

Joint goals and plans

- Problem: agents A and B need to move object X from room R1 to room R2
- Plan: A moves X from R1 to R2; B does nothing
- Plan': B moves X from R1 to R2; A does nothing

coordination of plans necessary

- What if we use stochastic policies instead of plans?

Multi-agent decision making

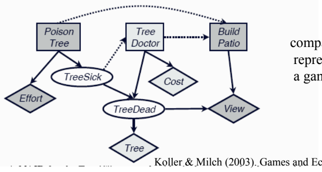
- … in cooperative, sequential games
- What happens to the MDP?
  - local views onto the world lead to partial observability
  - coordination requires extra actions to communicate
- POMDP formulation
  - MDP + observation function + belief state

- Multi-agent POMDP problems
  - potential for infinitely nested beliefs
  - potential for a non-stationary world (when learning)

Multi-agent graphical models

Example 3.1. Alice is considering building a patio behind her house, and the patio would be more valuable to her if she could get a clear view of the ocean. Unfortunately, there is a tree in her neighbor Bob's yard that blocks her view. Being somewhat unscrupulous, Alice considers poisoning Bob's tree, which would cost her some effort but might cause the tree to become sick. Bob cannot tell whether Alice has poisoned his tree, but he can tell if the tree is getting sick, and he has the option of calling in a tree doctor (at some cost). The attention of a tree doctor reduces the chance that the tree will die during the coming winter. Meanwhile, Alice must make a decision about building her patio before the weather gets too cold. When she makes this decision, she knows whether a tree doctor has come, but she cannot observe the health of the tree directly. A MAID for this scenario is shown in Fig. 3.

Alice: dark gray
Bob: light gray



compact
representation of
a game

Koller & Milch (2003). Games and Economic Behavior 45:1

Summary

- Game theory describes rational play in competitive (one-move) games with simultaneous moves
- Solutions to games are Nash equilibria where no agent gains by unilaterally changing its strategy
- Prisoner's Dilemma is a game where the dominant strategy leads to equilibrium with worse payoffs
- Mechanism design is inverse game theory (design rule system to maximize social welfare)
- Sequential cooperative games can be formulated as multi-agent partially observable MDPs