IntelligentAgent

Liangwei CHEN

November 2019

1 Real-world Games

- Mixed Nash Equilibrium may not be optimal for all participants implies Need for mediator for cooperation.

	Mediator	cooperate	defect
Mediator	(9, 9)	(10, 0)	(5, 5)
cooperate	(0, 10)	(9, 9)	(0, 10)
defect	(5, 5)	(10, 0)	(5, 5)

- Equilibrium types:
  - For simplicity we denotes σ as the equilibrium strategy distribution in all following expression
  - Pure Nash Equilibrium:

    $$s = \sigma,\ \forall i \in AGENTS, \forall s^{'} \in S_i,\ C_i(s) \leq C_i(s^{'}, s_{-i})$$

    Properties: May not exist, Best Price of Anarchy bound.
  - Mixed Nash Equilibrium:

    $$\forall i \in AGENTS, \forall s^{'} \in S_i,\ \mathbb{E}_{s\sim\sigma}[C_i(s)] \leq \mathbb{E}_{s\sim\sigma}[C_i(s^{'}, s_{-i})]$$

    Assume INDEPENDENT strategy decision. Existence promise. Hard to find.
  - Correlated Equilibrium:

    $$all i \in AGENTS,\ \forall s_i \in S_i\ s.t.\ \mathbb{P}_\sigma(s_i) > 0,$$

    $$\mathbb{E}_{s_{-i}\sim\sigma|s_i}[C_i(s_i, s_{-i})|s_i] \leq \mathbb{E}_{s_{-i}\sim\sigma|s_i}[C_i(s^{'}, s_{-i})|s_i]$$

    Interpretation: For any agent and any non-trivial deterministic strategy of this agent, if other agents choose their strategy according to correlated equilibrium distribution under the condition of this agent's strategy, then moving to other deterministic strategy cannot reduce the cost of this agent.
    The fact that MNE is CE can be seen by analyzing the optimality of each non-trivial deterministic strategy. The only difference between MNE and CE is that CE allows CORRELATED strategy distribution.

- Coarse Correlated Equilibrium:

$$\forall i \in AGENTS,\ \mathbb{E}_{s\sim\sigma}[C_i(s_i, s_{-i})] \leq \mathbb{E}_{s\sim\sigma}[C_i(s^{'}, s_{-i})]$$

  The only difference between CE and CCE is that CE promises optimality under every condition while CCE only promise optimality when consider all non-trivial strategies together.

- Relationship:
$$PNE \subset MNE \subset CE \subset CCE$$

- Negotiation
  - Alternating negotiation without discount factor: Agents propose in turn while the resource's value does not decrease with time. Only the last proposal matters.
  - Alternating negotiation with discount factor: Agents propose in turn while the resource's value decrease with time. Solve the trivial-but-last problem. Results depend on who start first.
  - Parallel negotiation: Agents exchange proposals in parallel. Solve the order problems above.
    Goal: Nash Bargaining Solution: Find the deal D^* such that:

    $$D^* = argmax_D(\Pi_{i\in AGENTS} r_i(D))$$

    Namely find the deal maximize the product of rewards.
    Algorithm: Monotonic concession protocol (2-agents).
    Since the reward function r is private to every agent, the agent needs to calculate a number containing its response to both proposals locally. These numbers should then be exchanged and indicate which of the proposal leads to bigger product of reward. The agents will then agree to move in that direction.
    Assume two agents i, j in the system. Agent i calculates $f_i = 1 - \frac{r_i(D_j)-r_i(D_{conflict})}{r_i(D_i)-r_i(D_{conflict})}$. Similarly for agent j.
    Notice that $f_i > 1\ iff\ r_i(D_j) < r_i(D_{conflict})$, in that case agent i would rather stop the negotiation and choose the conflicting plan directly, for instance in the scenario of wireless transmitting, it will attempt to transmit all the time regardless of what its peer does.
    Furthermore, $f_i > f_j$ indicates

    $$(r_i(D_i) - r_i(D_{conflict}))(r_j(D_i) - r_j(D_{conflict})) >$$

    $$(r_i(D_j) - r_i(D_{conflict}))(r_j(D_j) - r_j(D_{conflict}))$$

    . Thus the agent j will know he should adjust the plan to move in the direction of i's proposal.

2 Auction and Mechanism

- Auctions
  - Open-Box; First price: Bidding is public, the price paid finally is the highest bid.
    Implementation: Seller decrease the price till some buyer bid for it.
    Property: Decide at $v(2nd)+\epsilon$; Buyer need to speculate about others' values to decide the epsilon and the second highest value; Not suffer from collusion (the guy value it the most cannot collude with others to reduce the price)
  - Open-Box; Second price: Bidding is public, the price paid finally is the second highest bid
    Implementation: Increase bid each round till nobody is willing to bid higher
    Property: Decide at $v(2nd)$; Buyer don't need to speculate about others' values;Suffer from collusion; Incentive competitive
  - Closed-Box; First Price: Bidding is closed, the price decided finally is the highest bid
    Property: Single round; Decide at $v(2nd) + \epsilon$; Not suffer from collusion; Need speculation
  - Closed-Box; Second Price: Bidding is closed, the price decided finally is the second highest bid
    Property: Single round; Decide at $v(2nd)$; Suffer from collusion; No need for speculation; Truthfulness
  - Truthfulness: The property that an agent taking its true value as proposal.
  - Multi unit auction: The auctions with multiple same objects to be sold
    Problem: The buyer values the object more buy first with higher price. In other words, the price decreases along the auction.
    Solution: Decide at $(N + 1)$th price assuming N objects to be sold. This is the extension of Second-Price scheme.
  - Double auction: Multi buyers want to buy same object from multi seller. This is the abstraction of stock market.
    Implementation: Assume M sellers and N buyers, set the price to be arbitrary value between Mth and $M + 1$th highest bid among sellers and buyers.
    Property: $price = Mth\ highest$ seller incentive compatible; $price = M + 1th\ highest$ buyer incentive compatible
- Mechanism

- Definition: Mechanism is a game such that the dominant strategy or the equilibrium of the game leads to Social Choice Function's outcome.
- Property 1: Not all Social Choice Functions can be represented as a mechanism
- Property 2: Maximizing sum of declared valuation can be formulated as a mechanism (VCG tax mechanism)
  Let
  $$p_i(O) = tax_i(O) = \Sigma_{j \neq i} v_j(O) - v_j(O_{-i})$$
  such that O_{-i} is the optimal outcome without participation of i.
  Then by setting the reward function of each agent to be
  $$r_i(O) = v_i(O) - tax_i(O) = \Sigma_k v_i(O) - \Sigma_{j \neq i} v_i(O_{-i})$$
  One may see that

  * Optimizing reward function individually actually corresponds to optimizing the sum of declared valuation since the second term in the right hand side of the second equality is not related to self.
  * Agent obtain better result by participating in the game than not doing so even after paying the tax (Hint: compare $r_i(O)$ *to* $v_i(O_{-i})$)
  * The agent cannot benefit from underdeclaring or overdeclaring.