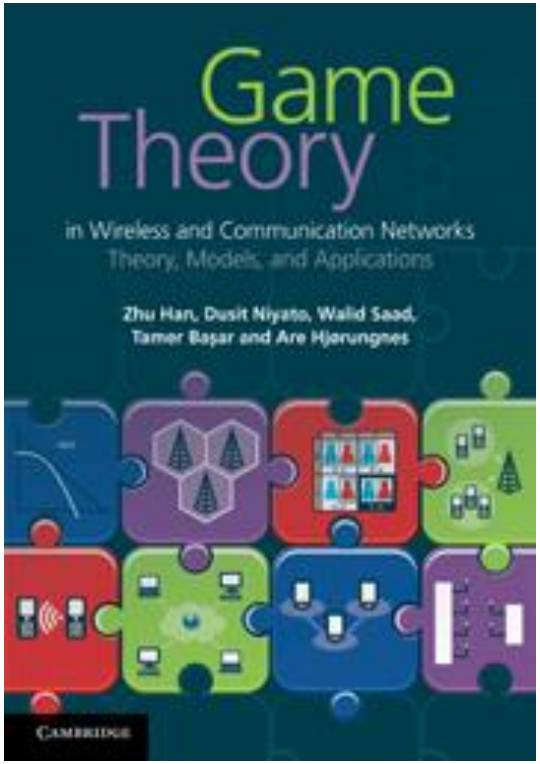
Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 2

Bayesian Games

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of Lecture Notes

- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, book 2
- Contract Theory, Lecture 8, book 2
- Stochastic Game, Lecture 9, book 2
- Learning in Game, Lecture 10, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 11, book 2
- Mean Field Game, Lecture 12, book 2
- Zero Determinant Strategy, Lecture 13, book 2
- Network Economy, Lecture 14, book 2
- Game in Machine Learning, Lecture 15, book 2

Overview of Bayesian Game

- Static Bayesian Games
- Bayesian Dynamic Games in Extensive Form
- Detailed Example: Cournot Duopoly Model with Incomplete Information
- Applications in Wireless Networks
  - Packet Forwarding
  - K-Player Bayesian Water-filling
  - Channel Access
  - Bandwidth Auction
- Summary

What is Bayesian Game?

Game in Strategic Form

- ***Complete Information:*** Each player has complete information regarding the elements of the game
- ***Dominant Strategy:*** Iterated deletion of other strategies
- ***Nash Equilibrium:*** Solution of the game in strategic form

Bayesian Game

- A game with incomplete information
- Each player has initial private information, **type**
- ***Bayesian Equilibrium:*** Solution of the Bayesian game

Bayesian Game

Definition

A Bayesian game is a strategic form game with incomplete information. It consists of:

- A set of **players**, N={1, …, n}; for each $i \in$ N,
- An **action** set, $A_i, (A = \times_{i \in N} A_i)$
- A **type** set, $\Theta_i, (\Theta = \times_{i \in N} \Theta_i)$
- A **probability** function, $p_i : \Theta_i \rightarrow \Delta(\Theta_{-i})$
- A **payoff** function, $u_i : A \times \Theta \rightarrow \mathbf{R}$

- The function p_i is player *i*'s **belief** about the types of the other players
- Payoff of player *i*, u_i is defined as a function of action, A and type, Θ
- Belief formulation captures the **uncertainty on the payoffs** of other players

Bayesian Game

Definition

Bayesian game $(N, \{A_i\}_{i\in N}, \{\Theta_i\}_{i\in N}, \{p_i\}_{i\in N}, \{u_i\}_{i\in N})$ is finite if N, A_i, and Θ_i are all finite

Definition (Pure strategy, Mixed strategy)

Given a Bayesian Game $(N, \{A_i\}_{i\in N}, \{\Theta_i\}_{i\in N}, \{p_i\}_{i\in N}, \{u_i\}_{i\in N})$, a pure strategy for player i is a function which maps player i's type into its action set

$$a_i : \Theta_i \to A_i$$

A mixed strategy for player i is

$$\alpha_i : \Theta_i \to \Delta(A_i) : \theta_i \to \alpha_i(.\,|\,\theta_i)$$

Bayesian Equilibrium

Definition

A Bayesian equilibrium of a Bayesian game is a mixed strategy profile $\alpha = (\alpha_i)_{i \in N}$, such that for every player $i \in$ N and every type $\theta_i \in \Theta_i$, we have

$$\alpha_i(.|\theta_i) \in \arg\max_{\gamma \in \Delta(A_i)} \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i}|\theta_i) \sum_{a \in A} \{ \prod_{j \in N \setminus \{i\}} \alpha_j(a_j|\theta_j) \} \gamma(a_i) u_i(a,\theta)$$

- Bayesian equilibrium is one of the mixed strategy profiles which maximizes each players' expected payoff for each type.

- This equilibrium is the solution for the Bayesian game. This equilibrium is the best response to each player's belief about the other player's mixed strategy.

Bayesian Dynamic Game

- Bayesian equilibrium results in complicated equilibria in dynamic game
- Refinement schemes do not always work in incomplete information
- Hence, perfect Bayesian equilibrium (demands optimal subsequent play)

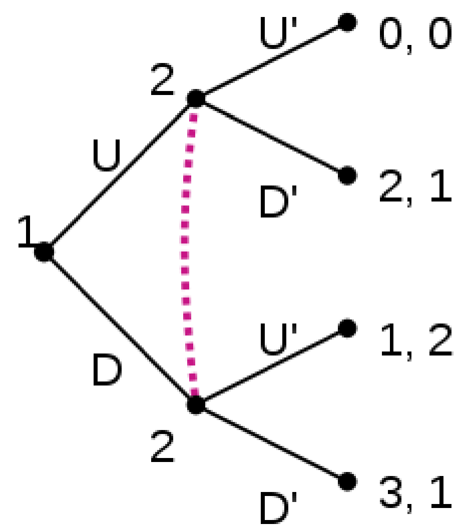
Definition 30 *A perfect Bayesian equilibrium is a strategy profile and a set of beliefs for each player such that:*

1. *at every information set, player i's strategy maximizes its payoff, given strategies of all other players, and player i's beliefs.*
2. *at information sets reached with positive probability when PBE strategy is played, beliefs are formed according to strategy and Bayes' rule when necessary.*
3. *at information sets that are reached with probability zero when PBE strategy is played, beliefs may be arbitrary but must be formed according to Bayes' rule when possible.*

Simple Example

- **Information is imperfect** since *player 2* does not know what *player 1* does when he comes to play.
- If both players are rational and both know that too, play in the game will be as follows according to perfect Bayesian equilibrium:

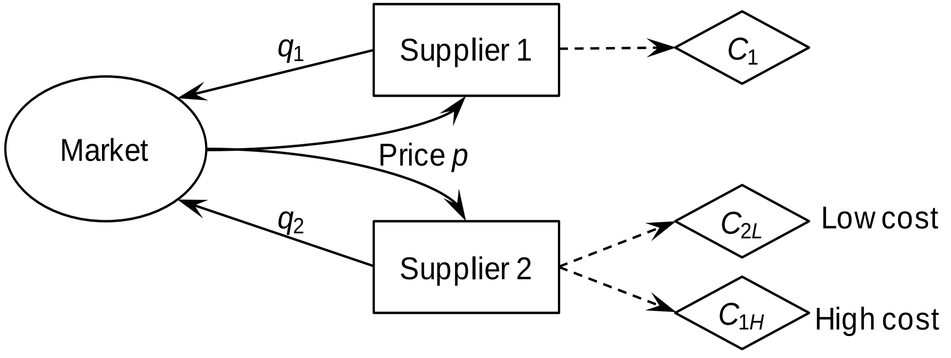
- *Player 1* likes to fool *player 2* into thinking he has played *U*, when he has actually played *D*, so that *player 2* will play *D'* and *player 1* will receive 3.
- In fact, there is a perfect Bayesian equilibrium where *player 1* plays *D* and *player 2* plays *U'* and *player 2* holds the belief that *player 1* will definitely play *D* (i.e *player 2* places a probability of 1 on the node reached if *player 1* plays *D*).
- In this equilibrium, every strategy is rational given the beliefs held and every belief is consistent with the strategies played. In this case, the perfect Bayesian equilibrium is the **only Nash equilibrium**.



Player 2 cannot observe *player 1*'s move

Example: Cournot Duopoly

Cournot duopoly model



(1) Players (2 firms): $N = \{1, 2\}$

(2) Action set (outcome of firms): $q_i \in \mathbf{R}_+, (i = 1, 2)$

(3) Type set: $\theta_1 = \{1\}, \theta_2 = \{3/4, 5/4\}$

(4) Probability function:

$$p(\theta_2 = 3/4 \mid \theta_1) = 1/2,\ p(\theta_2 = 5/4 \mid \theta_1) = 1/2$$

(5) Profit function:

$$u_1(q_1, q_2, \theta_1, \theta_2) = q_1(\theta_1 - q_1 - q_2)$$

$$u_2(q_1, q_2, \theta_1, \theta_2) = q_2(\theta_2 - q_1 - q_2)$$

Example: Cournot Duopoly

Bayesian equilibrium for pure strategy:

- The Bayesian equilibrium is a maximal point of expected payoff of firm 2, EP_2:

$$EP_2 = u_2 \qquad \frac{\partial EP_2}{\partial q_2}(q_1^*, q_2^*) = \theta_2 - q_1^* - 2q_2^* = 0$$

$$q_2^*(\theta_2) = (\theta_2 - q_1^*)/2, \ (\theta_2 = 3/4, 5/4)$$

- The expected payoff of firm 1, EP_1, is given as follows:

$$EP_1 = \frac{1}{2} q_1(\theta_1 - q_1 - q_2(3/4)) + \frac{1}{2} q_1(\theta_1 - q_1 - q_2(5/4))$$

Example: Cournot Duopoly

- Bayesian equilibrium is also the maximal point of expected payoff, EP_1:

$$\frac{\partial EP_1}{\partial q_1}(q_1^*, q_2^*) = 1 - 2q_1^* - \frac{1}{2}\{q_2^*(3/4) + q_2^*(5/4)\} = 0$$

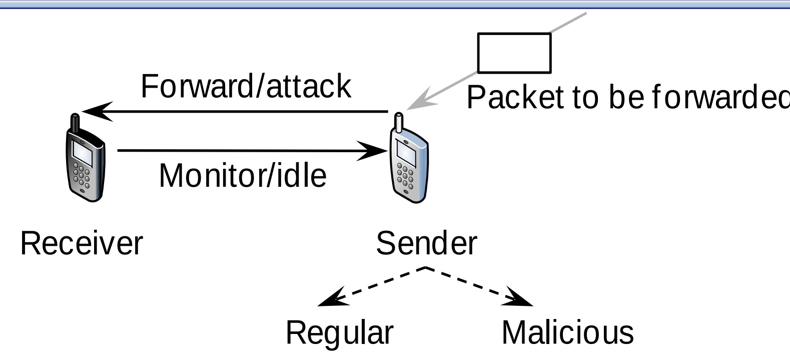
$$q_1^* = \frac{2 - q_2^*(3/4) - q_2^*(5/4)}{4}$$

- Solving the above equations, we can get Bayesian equilibrium as follows:

$$q_1^* = \frac{1}{3},\ q_2^*(3/4) = \frac{11}{24},\ q_2^*(5/4) = \frac{5}{24}.$$

Example: Packet Forwarding

- System model; Pure strategy
  - ***Sender type:*** Malicious or regular
  - ***Payoff:*** Malicious vs. regular



	Monitor	Stay idle
Attack	$(-G_A - C_A, G_A - C_M)$	$(G_A - C_A, -G_A)$
Forward	$(-C_F, -C_M)$	$(-C_F, 0)$

	Monitor	Stay idle
Forward	$(-C_F, -C_M)$	$(-C_F, 0)$

  - *Malicious sender:* $\mathbf{C_A}$- Cost of attack, $\mathbf{G_A}$- Attack success
  - *Regular sender:* $\mathbf{C_F}$- Cost of forwarding
  - *Receiver:* $\mathbf{C_M}$- Cost of monitoring, $\mathbf{-G_A}$- Cost of being attacked

- Mixed strategy
  - *Malicious sender:* Attack with a certain **probability**
  - *Regular sender:* Forward packet
  - *Receiver:* Monitor with a certain probability

Example: K-Player Bayesian Water-filling

- System model
  - R: Rate; P: Power; h: Channel gain

$$U_i = R_i = \log_2\left(1 + \frac{P_i(h_{i,i})h_{i,i}}{\sigma^2 + \sum_{j\neq i} P_j(h_{j,i})h_{j,i}}\right)$$

- A user knows the exact value of its own channel gain only
  - Channel gains of others are **random variables**, e.g. Shadow fading

$$L(R) = \overline{L}(R_{\min}) + 10n\log R - 10n\log R_{\min} + X_\sigma$$

- Path loss probability density function (belief of the other users)

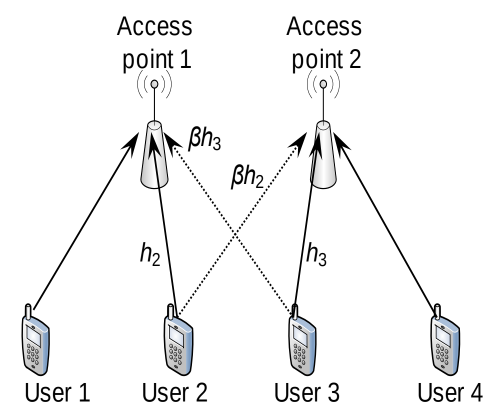
$$\alpha_L(l) = \begin{cases} \int_{a-C_0}^{l-C_0+3\sigma} g(x)\mathrm{d}x, & \text{if } a-3\sigma < l \leq a+3\sigma, \\ \int_{l-C_0-3\sigma}^{l-C_0+3\sigma} g(x)\mathrm{d}x, & \text{if } a+3\sigma < l \leq b-3\sigma, \\ \int_{l-C_0-3\sigma}^{b-C_0} g(x)\mathrm{d}x, & \text{if } b-3\sigma < l \leq b+3\sigma, \\ 0, & \text{otherwise,} \end{cases}$$

- The best response of user *i*

$$P_i^*(l_i) = \arg\max_{P_i} \int_{\mathcal{L}_{-i}} \alpha_L(\mathbf{l}_{-i}) U_i(P_i, \mathbf{p}_{-i}(\mathbf{l}_{-i}), (l_i, \mathbf{l}_{-i}))\mathrm{d}\mathbf{l}_{-i}$$

Example: Channel Access

- System model
  - Bayesian since do not know the other's channel (or type)
  - Transmission success $\gamma_i = \dfrac{Ph_i}{\beta \sum_{j \neq i} Ph_j + \sigma^2} \geq \gamma_{\text{thr}}$
  - Throughput $R_i(\mathbf{s}_{-i}) = \begin{cases} \log(1+\gamma_i), & \text{if } \gamma_i \geq \gamma_{\text{thr}}, \\ 0, & \text{otherwise,} \end{cases}$
  - Utility $U_i(s_i, \mathbf{s}_{-i}) = \begin{cases} 0, & \text{if } s_i = \text{backoff}, \\ R_i(\mathbf{s}_{-i}) - C, & \text{if } s_i = \text{transmit}, \end{cases}$
  - Bayesian NE $s_i^*(h_i) = \arg\max_{s_i \in \mathcal{S}_i} E\left(U(s_i, \mathbf{s}_{-i}(\mathbf{h}_{-i}), h_i)\right)$



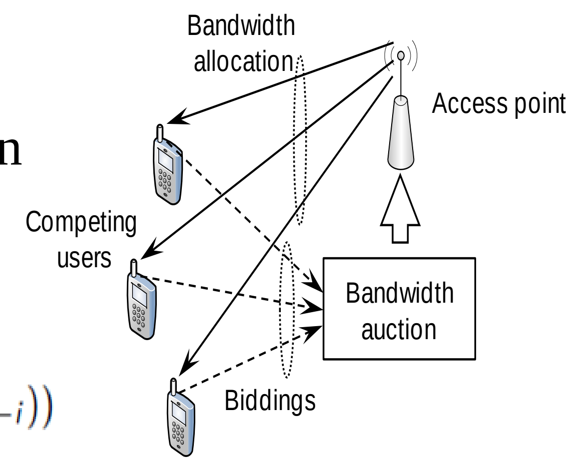
Interference among transmissions

- Threshold strategy
  - Interpretation of opportunistic spectrum access from the game theory point of view

$$s_i(h_i) = \begin{cases} \text{transmit}, & \text{if } h_i > h_{\text{thr},i}, \\ \text{backoff}, & \text{otherwise.} \end{cases}$$

Example: Bandwidth Auction

- System model
  - Allocated bandwidth $g_i = \dfrac{s_i t_i}{\sum_{j=1}^{N} s_j t_j} B,$
  - $\mathbf{s_i}$ - Bid, $\mathbf{B}$ - Total bandwidth, $\mathbf{t_i}$ - Time duration of connection, $\mathbf{U}$ - utility
  - Pdf of $\mathbf{t_i}$ of other nodes $\alpha_T(t) = \dfrac{1}{\beta}\exp(-t/\beta) > 0,$
  - Bayesian NE $s_i^*(t_i) \;=\; \arg\max_{s_i} E\left(\mathscr{U}_i(g_i, s_i, \mathbf{s}_{-i}, t_i, \mathbf{t}_{-i})\right)$



- Iterative algorithm to obtain Bayesian NE

Initialize iteration counter $k = 1$.
Access point receives $s_i[k]$, $t_i[k]$, and $g_{\mathrm{thr},i}$ from all vehicular nodes.
Access point computes $g_i[k-1]$ from (4.56) and allocates bandwidth to the vehicular nodes.
repeat
  $k \leftarrow k + 1$.
  $s_i^*[k] \leftarrow \arg\max_{s_i} E\left(\mathscr{U}_i(g_i[k-1], s_i[k-1], \mathbf{s}_{-i}[k-1], t_i[k-1], \mathbf{t}_{-i})\right)$.
  Vehicular node i sends $s_i^*[k]$ to access point.
  Access point computes $g_i[k]$ from (4.56) and allocates bandwidth to the vehicular nodes.
until $\max_i |s_i^*[k] - s_i^*[k-1]| \le \epsilon$.

[16]

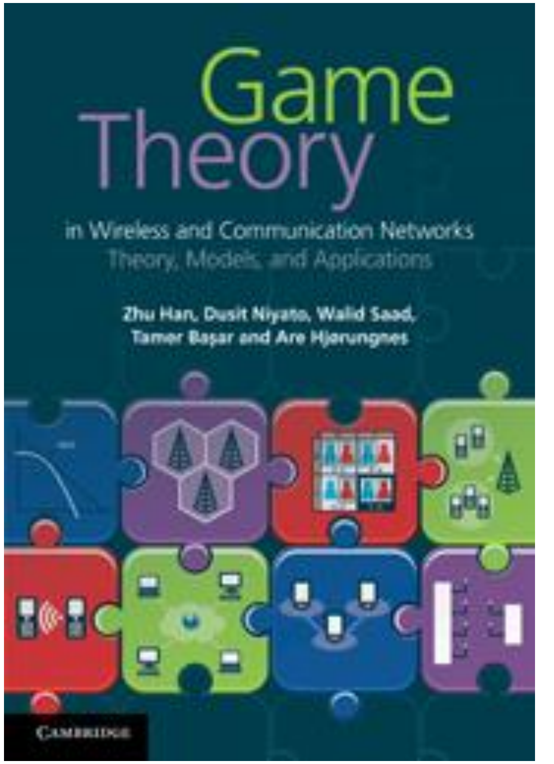
Summary

- Games with incomplete information (i.e., Bayesian game) can be used to analyze situations where a player does not know the preference (i.e., payoff) of his opponents.
- This is a common situation in wireless communications and networking where there is no centralized controller to maintain the information of all users. Also, the users may not reveal the private information to others.
- The details of the Bayesian game framework were studied.
- Some applications of the Bayesian game framework in wireless communications and networking were discussed.

Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 5
Cooperative Game

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of Lecture Notes

- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, Chapter 2, book 2
- Contract Theory, Lecture 8, Chapter 3, book 2
- Learning in Game, Lecture 9, Chapter 6, book 2
- Stochastic Game, Lecture 10, Chapter 4, book 2
- Game with Bounded Rationality, Lecture 11, Chapter 5, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 12, Chapter 7, book 2
- Zero Determinant Strategy, Lecture 13, Chapter 8, book 2
- Mean Field Game, Lecture 14, book 2
- Network Economy, Lecture 15, book 2

Introduction

- Introduction to cooperative game
- Bargaining solution
  - Nash Bargaining Solution
  - Kalai – Smorodinsky Bargaining Solution
  - Rubinstein Bargaining Process
  - Example of Bargaining in Wireless Networks
- Coalitional game
  - Class I: Canonical coalitional games
  - Class II: Coalition formation games
  - Class III: Coalitional graph games
- Summary

Cooperation in Wireless Networks

- Cooperation in wireless networks
  - Cooperation among network nodes
    - *Gains in terms of capacity, energy conservation or improved Bit Error Rate (BER)*
  - Ubiquitous in many networks
    - *Cognitive radio, sensor networks, WiMAX, ….*
- Levels of cooperation
  - Network Layer Cooperation
    - *Routing and Packet forwarding*
  - Physical Layer Cooperation
    - *Traditional Relay channel*
    - *Virtual MIMO*

Cooperative Game Theory

- Players have mutual benefit to cooperate
  - Startup company: everybody wants IPO, while competing for more stock shares.
  - Coalition in Parlement
- Namely two types
  - Bargaining problems
  - Coalitional game
- For coalitional game
  - Definition and key concepts
  - New classification
  - Applications in wireless networks

Walid Saad, Zhu Han, Merouane Debbah, Are Hjorungnes, and Tamer Basar, ``Coalitional Game Theory for Communication Networks'', IEEE Signal Processing Magazine, Special Issue on Game Theory, p.p. 77-97, September 2009.

Introduction to Bargaining

- Bargaining situation
  - A number of individuals have a **common** interest to cooperate but a conflicting interest on **how to cooperate**
- Key tradeoff
  - Players wish to reach an agreement rather than disagree but…
  - Each player is self interested
- What is bargaining?
  - Process through which the players on their own attempt to reach an agreement
  - Can be tedious, involving offers and counter-offers, negotiations, etc.
- Bargaining theory studies these situations, their outcome, and the bargaining process

Introduction

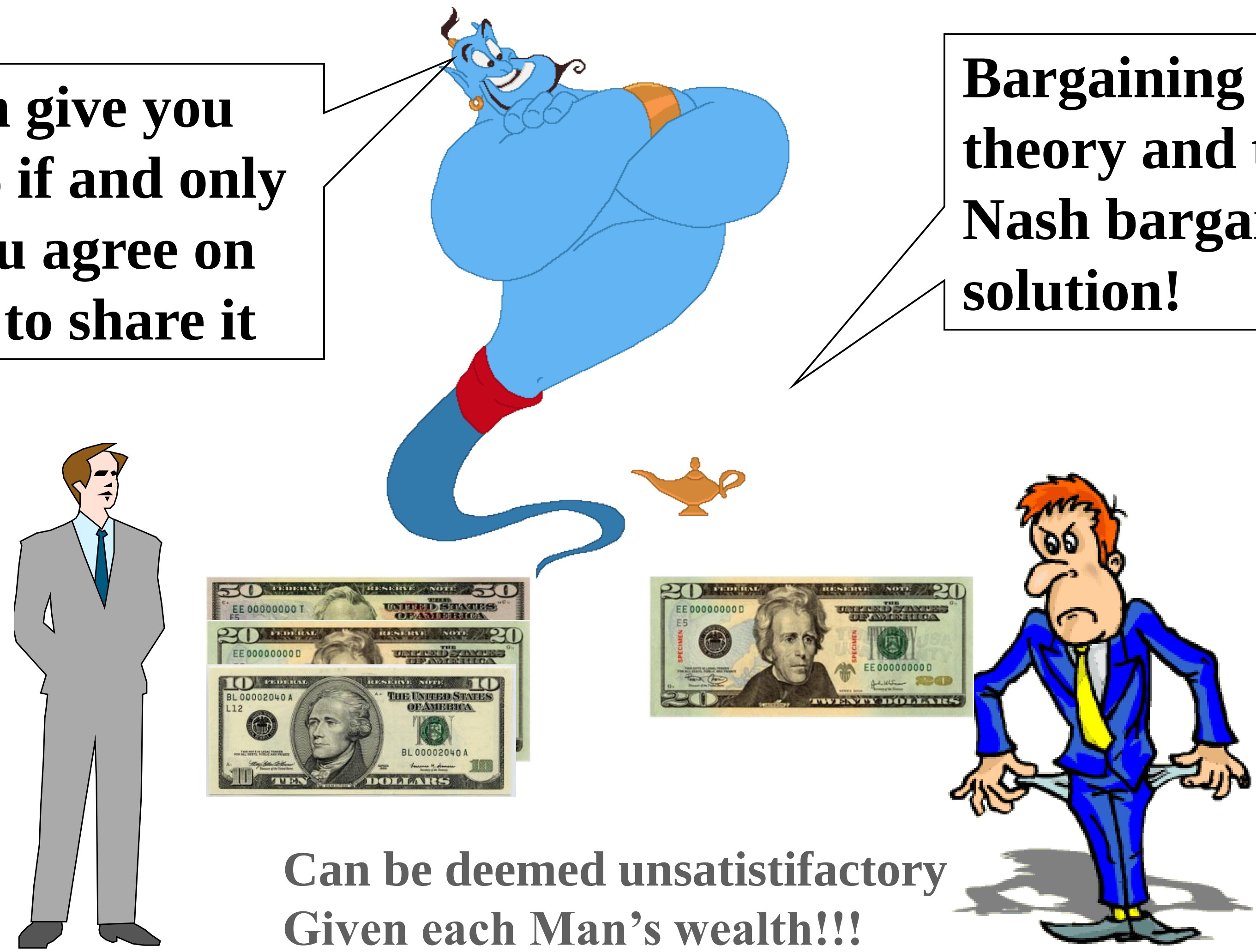
- Key issues in bargaining
  1. The players must inspect efficiency and the effect of delay and disagreement on it
     - *They seek a jointly efficient mutual agreement*
  2. Distribution of the gains from the agreement
     - *Which element from the efficient set must the players elect?*
  3. What are the joint strategies that the players must choose to get the desired outcome?
  4. How to finally enforce the agreement?

- Link to game theory
  - Issues 1 and 2 are tackled traditionally by cooperative game theory
  - Issues 3 and 4 are strongly linked to non-cooperative game theory

Motivating Example

I can give you 100$ if and only if you agree on how to share it

Bargaining theory and the Nash bargaining solution!



Can be deemed unsatistifactory
Given each Man's wealth!!!

Rich Man

The Nash Bargaining Solution

- **John Nash**'s approach
  - When presented with a bargaining problem such as the rich man – poor man case, how can we pick a reasonable outcome?
  - Interested in the **outcome** rather than the **process**
- **The Nash Bargaining Solution** was proposed in 1950 using an axiomatic approach and is considered as one of the key foundations of bargaining problems

The Nash Bargaining Solution

- Given a bargaining problem between **two** players
- Consider a **utility** region *S* that is compact and convex
  - A utility is a function that assigns a value to every player, given the strategy choices of **both** players
- Define the **disagreement or threat point** *d in S* which corresponds to the minimum utilities that the players want to achieve
- A Nash bargaining problem is defined by the pair *(S,d)*

The Nash Bargaining Solution

- Can we find a *bargaining solution*, i.e., a function f that specifies a **unique** outcome $f(S,d) \epsilon S$?
- Axiomatic approach proposed by Nash
  - Axiom 1: Feasibility
  - Axiom 2: Pareto efficiency
  - Axiom 3: Symmetry
  - Axiom 4: Invariance to linear transformation
  - Axiom 5: Independence of irrelevant alternatives

The Nash Bargaining Solution

- Axiom 1: Feasibility
  - Can be sometimes put as part of the definition of the space S
- Feasibility implies that
  - The outcome of the bargaining process, denoted (u^*,v^*) cannot be worse than the disagreement point $d = (d_1,d_2)$, i.e., $(u^*,v^*) \geq (d_1,d_2)$
  - Strict inequality is sometimes defined
- Trivial requirement but important: the disagreement point is a benchmark and its selection is very important in a problem!

The Nash Bargaining Solution

- Axiom 2: Pareto efficiency
  - Players need to do as well as they can without hurting one another
- At the bargaining outcome, no player can improve without decreasing the other player's utility
  - Pareto boundary of the utility region
- Formally, no point (u,v) ϵ S exists such that $u > u^*$ **and** $v \geq v^*$ or $u \geq u^*$ **and** $v > v^*$

The Nash Bargaining Solution

- Axiom 3: Symmetry
  - If the utility region is symmetric around a line with slope 45 degrees then the outcome will lie on the line of symmetry
  - Formally, if $d_1 = d_2$ and S is symmetric around $u = v$, then $u^* = v^*$
- Axiom 4: Invariance to linear transformation
  - Simple axiom stating that the bargaining outcome varies linearly if the utilities are scaled using an affine transformation

The Nash Bargaining Solution

- Axiom 5: Independence of irrelevant alternatives
  - If the solution of the bargaining problem lies in a subset *U* of *S*, then the outcome does not vary if bargaining is performed on *U* instead of *S*
- Controversial axiom
  - If we increase the maximum utility achievable by a player, the outcome does not change!
  - It is shown that although the bargaining power of one player might improve in the bigger set, the other would not
  - We will explore an alternative in later slides

The Nash Bargaining Solution

- Nash showed that there exists a unique solution *f* satisfying the axioms, and it takes the following form:

$$(u^*, v^*) = f(S, d) = \max_{(u,v)\in S} (u - d_1)(v - d_2)$$

When $d_1 = d_2 = 0$, this is equivalent to the famous solution of telecommunication networks: Proportional fairness

Known as the Nash product

Rich man – poor man problem revisited

- Considering logarithmic utilities and considering that what the men's current wealth is as the disagreement point
  - The Nash solution dictates that the rich man receives **a larger** share of the 100$
- Is it fair?
  - Fairness is subjective here, the rich man has more bargaining power so he can threaten more to stop the deal
    - *The poor man also values little money big as he is already poor!*
  - Variant of the problem considers the 100$ as a debt, and, in that case, the NBS becomes fair, the richer you are the more you pay!

The Nash Bargaining Solution

- The NBS is easily extended to the N-person case
  - The utility space becomes N-dimensional and the disagreement point as well
  - Computational complexity definitely increases and coordination on a larger scale is required
- Solution to the following maximization problem

$$(u_1^*, \ldots, u_N^*) = f(S, d) = \max_{(u_1, \ldots, u_N) \in S} \prod_{i=1}^{N} (u_i - d_i)$$

The Nash Bargaining Solution

- If we drop the Symmetry axiom we define the Generalize Nash Bargaining Solution
- Solution to the following maximization problem

$$(u_1^*, \ldots, u_N^*) = f(S, d) = \max_{(u_1, \ldots, u_N) \in S} \prod_{i=1}^{N} (u_i - d_i)^{\alpha_i}$$

Value between 0 and 1 representing the bargaining power of player *i*
If equal bargaining powers are used, this is equivalent to the NBS

Nash Bargaining Solution – Summary

- The NBS/GNBS are a very interesting concept for allocating utilities in a bargaining problem
  - Provide Pareto optimality
  - Account for the bargaining power of the players but..
  - Can be unfair, e.g., the rich man – poor man problem
  - Require convexity of the utility region
  - Independence of irrelevant alternatives axiom
  - Provide only a static solution for the problem, i.e., no discussion of the bargaining process
- Alternatives?
  - The Kalai – Smorodinsky solution
  - Dynamic bargaining and the Rubinstein process

Kalai – Smorodinsky Bargaining Solution

- Kalai and Smorodinsky (1974) proposed to replace the IIA axiom with an axiom of individual monotonicity
  - Expanding the utility space *S* in a direction favorable to a player *i* (forming *S'*) implies that the utility of player *i* obtained in *S'* is higher
- They showed that, for a 2-person game, a unique bargaining solution satisfying the individual monotonicity axiom + Nash axioms exists
  - Extended later to N-person game and asymmteric case

Kalai – Smorodinsky Bargaining Solution

- In a utility space S the Kalai – Smorodinsky solution is a unique point satisfying the following equation:

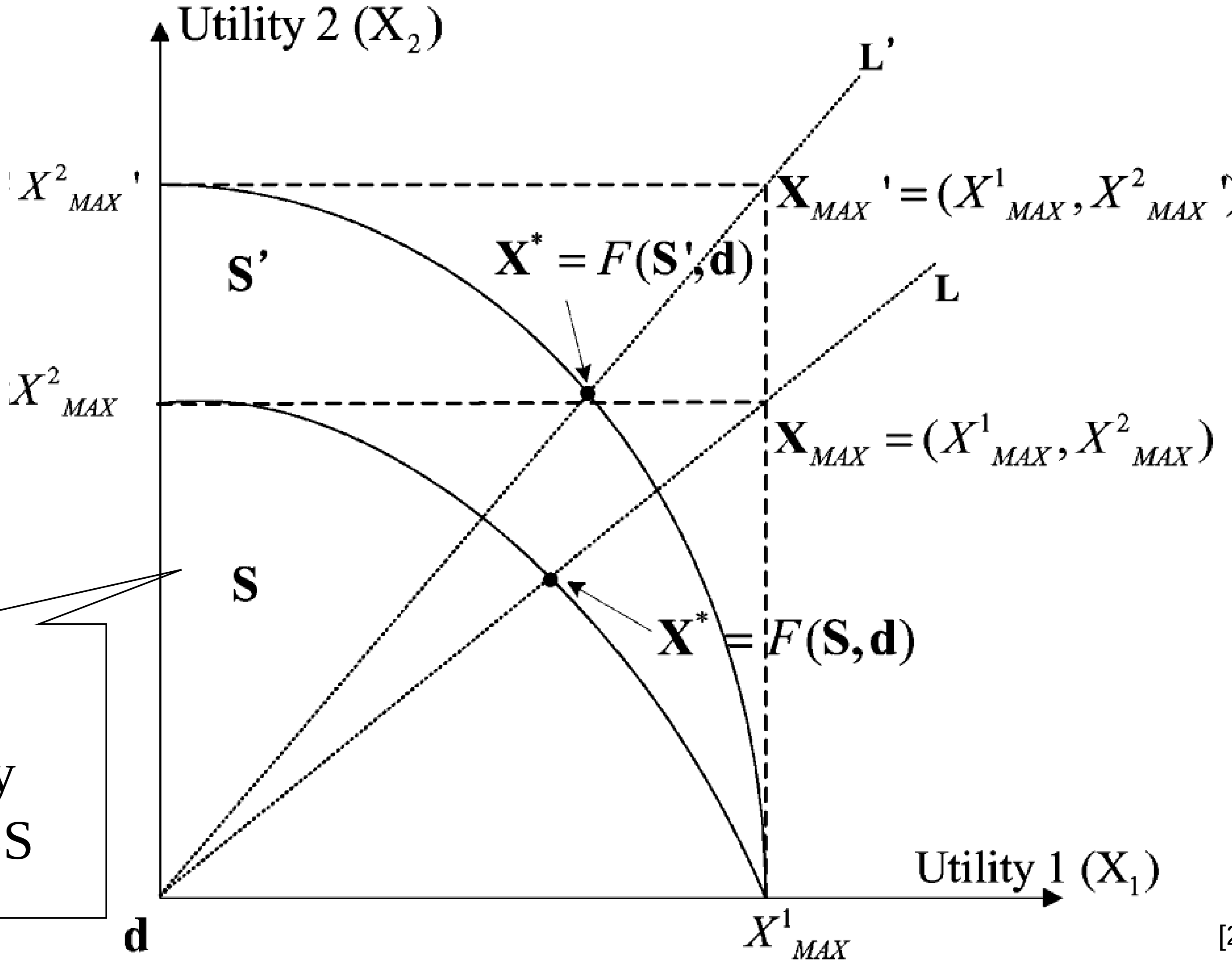
$$(u_1^*, \ldots, u_N^*) = (d_1, \ldots, d_N) + \lambda\left((i_1, \ldots, i_N) - (d_1, \ldots, d_N)\right)$$

Maximum λ such that the point d + λ (I – d) is still in S

Ideal point:
Point of best achievable utilities for the players

- The KSBS is simply the intersection between the Pareto boundary and the line connecting the disagreement point to the ideal point!

The KSBS - Illustration



Individual monotonicity and the KSBS

KSBS vs NBS

- In KSBS, the player having a bigger potential achievable utility receives more
- In NBS, the player with a bigger initial position, e.g., the rich man, gets a better share
- KSBS does not require convexity of the utility space
- Note on KSBS: it can also be used in a *generalized* sense with bargaining powers

Dynamic Bargaining

- The NBS and the KSBS are both static solutions in the sense that we only care about the outcome
  - How about the bargaining process?
- Dynamic bargaining
  - Interested in the players interactions to reach an agreement
  - Broader than static bargaining, although linked to it
  - In this trial lecture, we cover the Rubinstein process although many others exists

Rubinstein Bargaining Process

- Two players A and B bargain over the division of a cake of size 1

- Alternating-offers process
  - At time 0, A makes an offer to B
  - If B accepts, the game ends, otherwise
  - B rejects and makes a counter offer at time $\Delta > 0$
  - The process continues infinitely until an agreement is reached

- The payoff of a player i at any time is $x_i\delta^t$
  - x_i is the share of the cake for player i, $0 < \delta < 1$ a discount factor and t is the number of time intervals Δ elapsed
  - The discount factor is also function of a discount rate

Rubinstein Bargaining Process

- Rubinstein (1982) modeled this process as an extensive form non-cooperative game
  - When making an offer the player's strategy is the value of the share he requests
  - When responding to an offer the strategy is either accept or reject the offer (as a function of the history of the game so far)
- Rubinstein showed that the game admits a **unique subgame perfect equilibrium**
  - The equilibrium is also Pareto efficient

Rubinstein Bargaining Process

- At the equilibrium the shares of the players are

$$x_A^* = \frac{1}{1+\delta}, \quad x_B^* = \frac{\delta}{1+\delta}$$

  - At date t = 0, player *A* offers x_A, player *B* accepts and the game ends
  - First-mover advantage, player *A* gets more than player *B* but the result is Pareto efficient
  - As δ increases (the interval Δ decreases) this advantage starts to disappear
  - The rations depend highly on the relative discount rates
    - *The winner is the strongest*

Rubinstein Bargaining Process

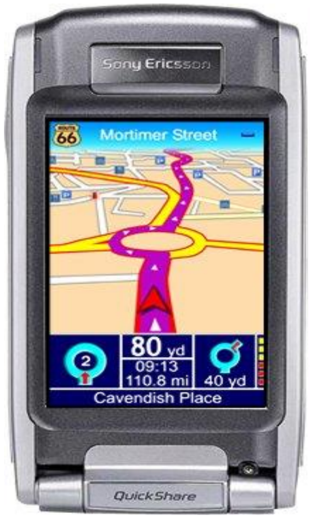
- The Rubinstein model shows that being more patient increases your bargaining power!
  - The smaller the cost of "haggling", the more waiting time you can sustain, the higher is your bargaining power
- If the process is frictionless then it becomes indeterminate!
- We reach NBS/GNBS as Δ goes to *0*

Other dynamic bargaining solutions

- The Rubinstein model is given for 2-player case, recent bargaining literature looks into variants for N-player games
- Varying preferences over time (non-stationary models)
- Emphasis on tools from non-cooperative games, e.g., repeated games, extensive form games, or stochastic games
- Dynamic bargaining is **hot** in bargaining literature nowadays

Bargaining in Wireless Networks

In Recent Years **Past Decade**



Voice based services

Video
Multimedia
Gaming
Instant Messaging & Presence
Data Transfer

Need for fair and efficient sharing of the scarce radio resources: frequency, power, etc.

Stringent QoS Requirements on resource constrainted networks!

Bargaining theory!

NBS in Gaussian Interference Channel (1)

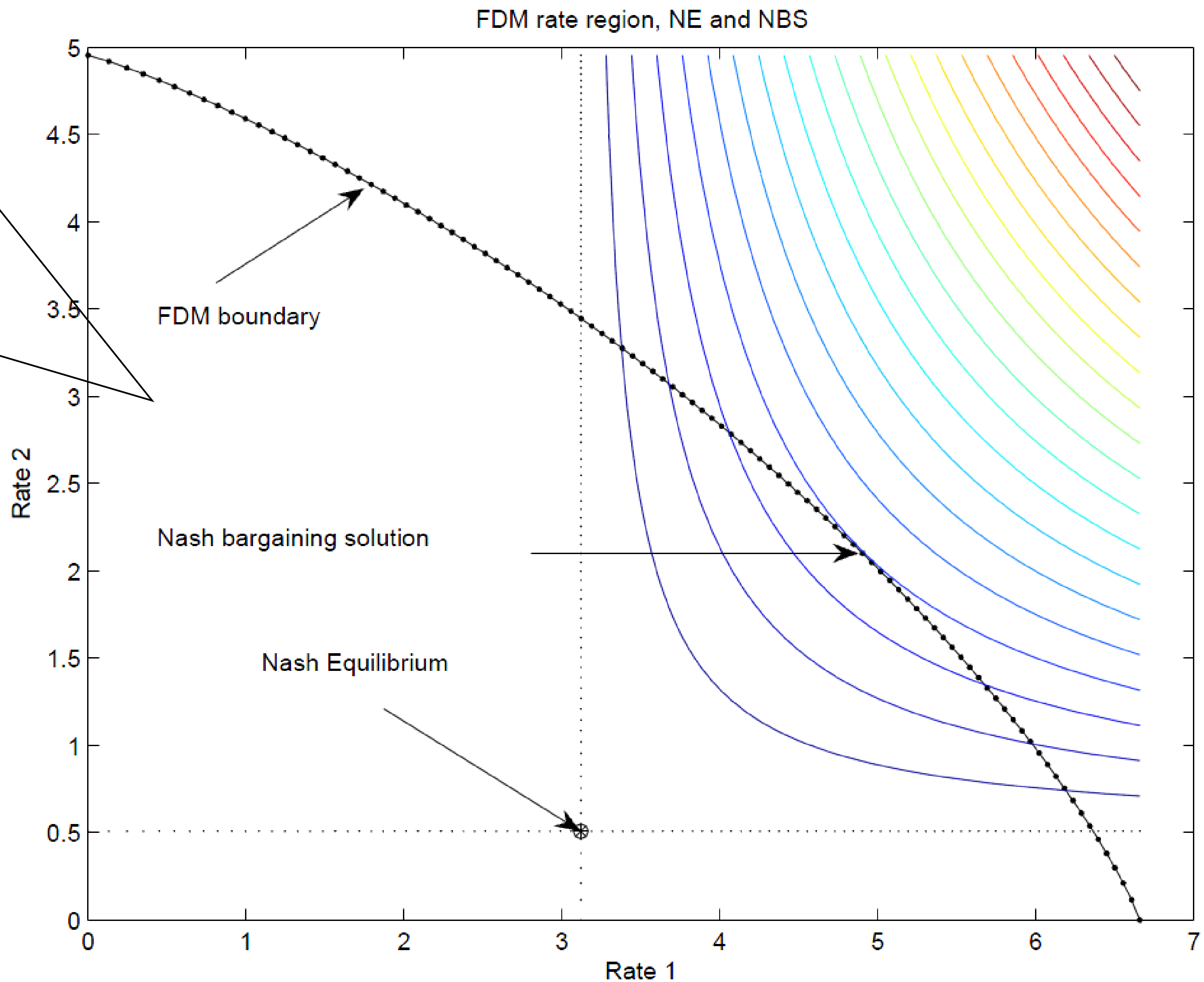
- A. Leshem and E. Zehavi, IEEE JSAC, vol. 26, no. 7, 2008
- Flat frequency channel
  - Nash equilibrium rates when each user transmits at its maximum power
  - Can we do better by cooperation, e.g., using FDM?
- A Nash Bargaining Problem
  - The network users are the players
  - The strategy of player n is the fraction ρ_n of the frequency band used by this player
  - The utility is simply the **rate**
  - What can we say about the NBS?

NBS in Gaussian Interference Channel (2)

- **Main result:** under certain conditions on the SNR of the users, the NBS exists
  - **Existence is in the sense that there exists a point better than the minimum point, i.e., the Nash equilibrium**
- What does it mean, really?
  - Using FDM, under certain SNR conditions (depending on channel gains), is better than acting non-cooperatively
  - A unique division of frequency exists that achieves the NBS
  - Applicable also to the case where the utility is a log of the rate

NBS in Gaussian Interference Channel (3)

Up to 1.6 and 4 times improvement for NBS over the NE for users 1 and 2, respectively



NBS in Gaussian Interference Channel – Frequency selective

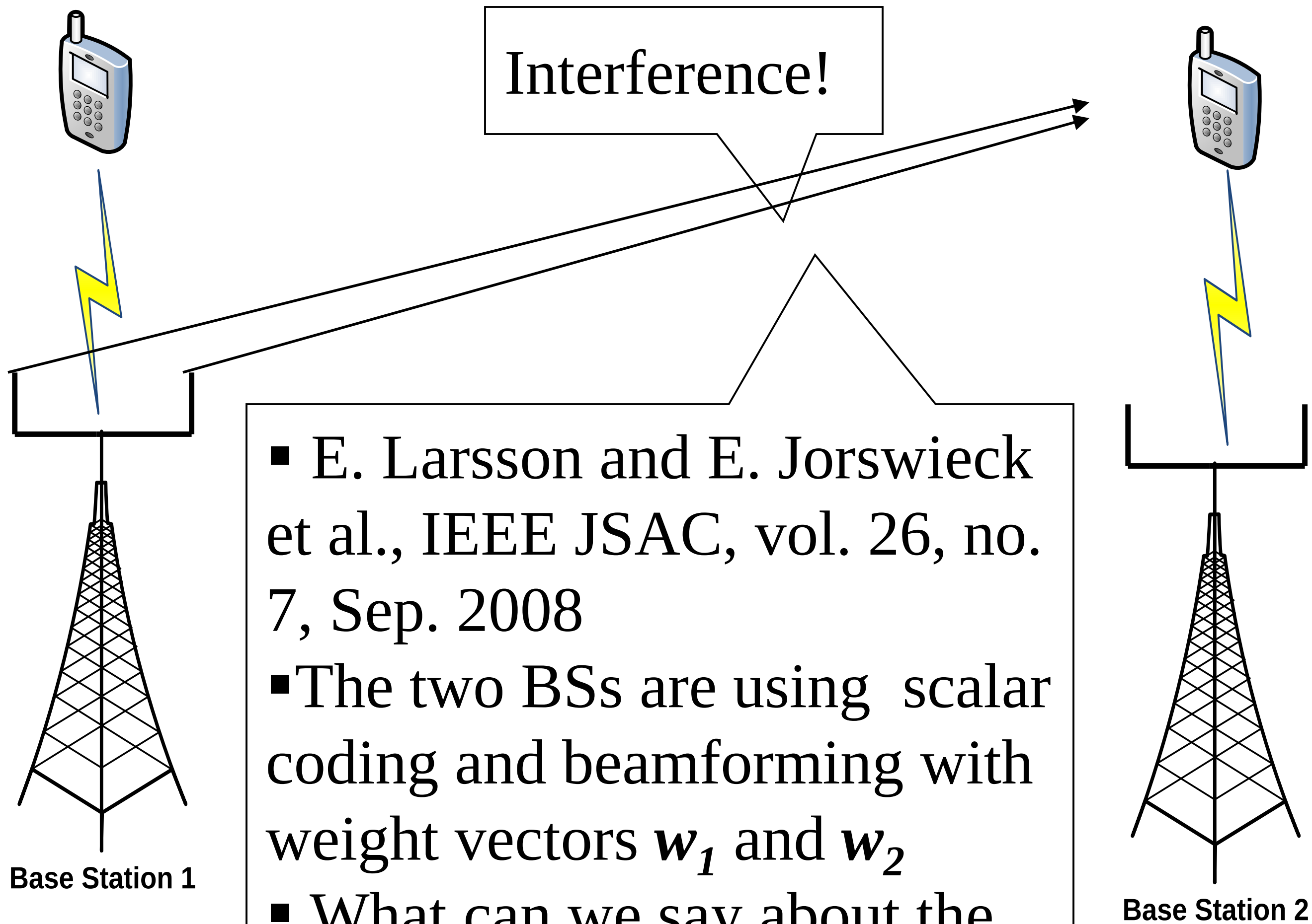
- What about the frequency selective case?
- Considered model
  - Joint FDM/TDM scheme whereby the players transmit over K frequency bins and each player i uses each frequency k for a certain fraction of time α_{ik}
  - The utility is the rate

$$R_i = \sum_{k=1}^{K} R_i(k) = \sum_{k=1}^{K} \alpha_{ik} \log_2 \left(1 + \frac{|h_{ii}(k)|^2 p_i(k)}{\sigma_i^2(k)} \right)$$

NBS in Gaussian Interference Channel – Frequency selective

- Main results
  - If the players have **different** rates ratios at each frequency, then, at the optimal NBS solution, at most $\binom{N}{2}$ frequencies are actually shared between the users
  - For 2 players, at most **one** frequency is shared between the two users
- Algorithms
  - Convex optimization can be used for finding the NBS of the N person interference channel game
  - For 2 players, an efficient $O(K \log K)$ algorithm is proposed

NBS in MISO Interference Channel



- E. Larsson and E. Jorswieck et al., IEEE JSAC, vol. 26, no. 7, Sep. 2008
- The two BSs are using scalar coding and beamforming with weight vectors $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$
- What can we say about the rates using game theory?

NBS in MISO Interference Channel

- The Nash Equilibrium point is unique and corresponds to the maximum-ratio transmission beamforming vectors

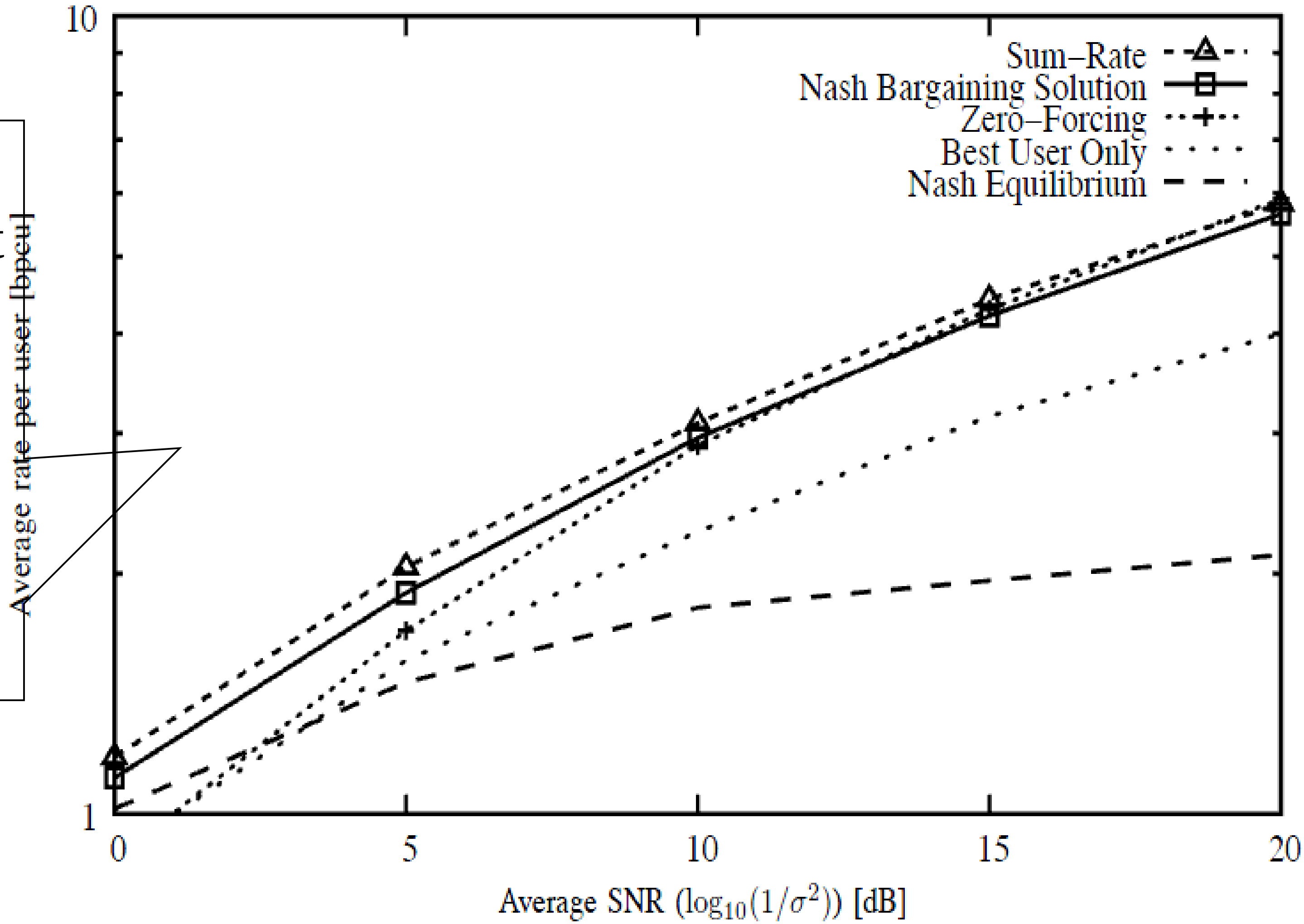
$$w_1^{\mathrm{NE}} = \frac{h_{11}^*}{\|h_{11}\|} \qquad \text{and} \qquad w_2^{\mathrm{NE}} = \frac{h_{22}^*}{\|h_{22}\|}$$

- At low SNR, the NE has an OK performance
- At high SNR, the NE is highly inefficient
- Any better alternative?

NBS in MISO Interference Channel

- The Nash Bargaining solution
  - Players are the two BSs
  - The utility is the rate
  - Disagreement point is the NE
- The Nash Bargaining solution achieves a Pareto optimal rate vector for the BSs better than the NE
  - Can be found graphically (intersection of the Pareto boundary with the contours of the Nash function) but..
  - Implicitly the BSs need to negotiate over some side channel
  - The rate region must be extended to become **convex** (using the convex hull/time share)

NBS in MISO Interference Channel



NBS almost as good as the sum-rate maximizing point!

Multimedia Resource Management through Bargaining

- H. Park and M. van der Schaar, IEEE TSP, vol. 55 No. 7, Jul. 2007
- Consider *n* video transmitters seeking to share a wireless or wired network's bandwidth
- Utility is a form of the PSNR (without the log), defined as follows

$$U_i(x_i) \triangleq \frac{c}{D_i} = \frac{c \cdot (x_i - R_{0i})}{D_{0i}(x_i - R_{0i}) + \mu_i}$$

Constants depending on the video characteristics

Distortion

Constant

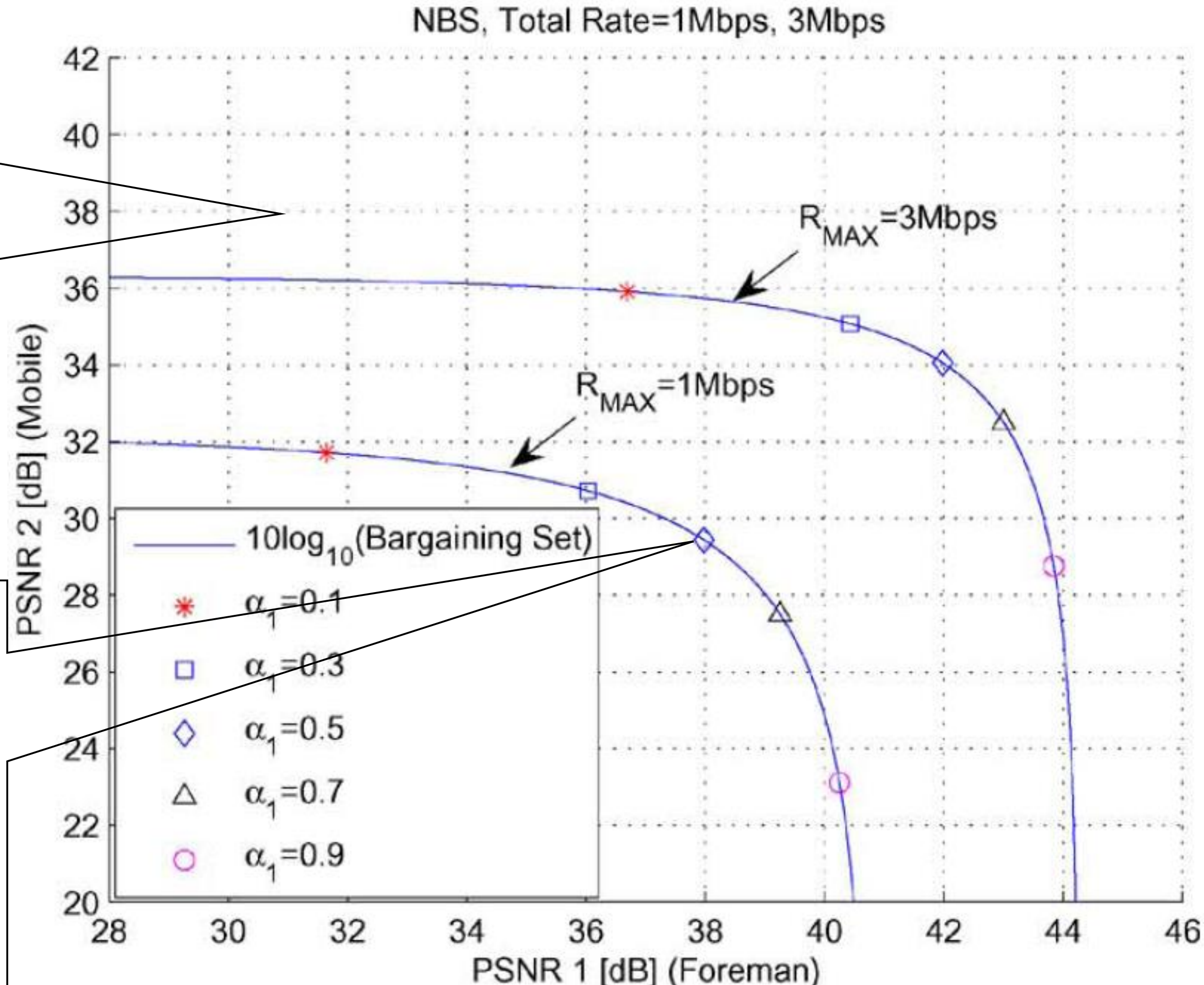
Allocated rate to user *i*

Multimedia Resource Management through Bargaining

- Key question: how to allocate the rates x_i taking into account
  - Optimality
  - Fairness
- Two different approaches
  - Generalized Nash Bargaining Solution
  - Kalai - Smorodinsky Bargaining Solution

Multimedia Resource Management through Bargaining - GNBS

As the bargaining power increases the PSNR increases



Adjusting the Bargaining Power to provide quality for both users

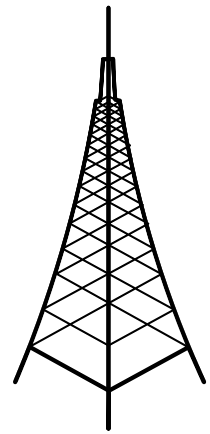
Bargaining in Cognitive Radio Networks – Spectrum Sensing

- Secondary users in a cognitive network need to sense the spectrum in order to decide on whether to access it or not
- Two important problems of spectrum sensing
  - Missing the detection of the primary user
  - Time spent for sensing
- Can we use dynamic bargaining to tackle these problems?
  - M. Pan and Y. Fang, MILCOM 2008

Bargaining in Cognitive Radio Networks – Spectrum Sensing

I can help you but I cannot spend too much in sensing although it is better for both of us

Secondary User



Primary User (Licensed user)

I can help you but I cannot spend too much in sensing although it is better for both of us

Secondary User

Bargaining situation!
Two players want to cooperate but they are self-interested

Bargaining in Cognitive Radio Networks – Spectrum Sensing

- Two secondary users bargain over sharing a time period
  - When and how long does each player spend in sensing given the loss in transmission time?
- Rubinstein problem of alternating offers
  - Two players
  - The share of the cake is the share of the time period each player uses for sensing
  - The utilities are a function of the time share, the rate, and the probability of detection
  - Discount rate is function of the PU – SU distance

Bargaining in Cognitive Radio Networks – Spectrum Sensing

- It is shown that the process maps to the Rubinstein solution
- The first mover advantage is decided based on the primary user behavior
  - When the PU is on, the user closest to it is the first mover
  - When the PU is off, the user farthest to it is the first mover
  - Expected utilities given the PU behavior
- Drawbacks
  - Process is limited to 2-player scenario
  - Limited mobility
  - Too much reliance on the distance to the PU!

Other applications of Bargaining Theory

- **MIMO, Information Theory, and Interference channel**
  - Channel allocation in OFDMA networks using NBS (Z. Han et al., IEEE TCOM, vol. 53, no. 8, Aug. 2005)
  - Survey of bargaining in OFDMA networks (A. Ibing and H. Boche, Asilomar 2007)
  - Similar to the MISO channel discussed in this lecture, NBS can be used in SIMO or MIMO settings (see E. Larsson et al., IEEE SPM, Special issue on Game Theory, Sept. 2009)
  - Extension of NBS to cover log-convex sets in a scenario where the utility is the SIR (H. Boche et al., IEEE/ACM TCN, vol. 17, no. 5, Oct. 2009)
  - NBS can be used to allocate the rates in an interference channel when the receivers **cooperate** (S. Mathur et al., IEEE JSAC, Sep. 2008)
  - KSBS in MIMO interference networks, (M. Nokleby and L. Swindlehurst, IEEE ICCCN, Aug. 2008)
  - NBS and MIMO interference systems, (Z. Chen et al., IEEE ICC, 2009)
  - NBS and frequency selective channels with precoding, (J. Gao et al., IEEE TSP, to appear, 2010)

Other applications of Bargaining Theory

- **Cognitive radio**
  - Spectrum allocation using local bargaining (L. Cao and H. Zheng, IEEE SECON, 2005)
  - Dynamic spectrum allocation using NBS and dynamic games (Z. Ji and K. J. Liu, IEEE JSAC, vol. 26, no. 1, Jan. 2008)
  - Channel and power assignment in cognitive networks (Attar et al., IEEE TWC, vol. 8, no. 4, Apr. 2009)
  - Distributed spectrum sharing through bargaining (Komali et al., IEEE TWC, vol. 8, no. 10, Oct. 2009)
  - Several papers by Hosseinabadi from the group of Jean-Pierre Hubaux at EPFL using Nash Bargaining in resource allocation for cognitive radio
- **Others**
  - KSBS for common radio resource management (M. Kahn et al., Gamenets, May 2009)
  - Using Nash bargaining for improving wireless access in **vehicular networks** (B. Shrestha et al., IEEE GLOBECOM, 2008)

Bargaining Theory in Wireless Networks – Remarks and Conclusions

- Most of the work on bargaining is focused on rate or channel allocation, MIMO, and interference channel
  - Due to the correlation of the rate region with the bargaining region
- Nash bargaining is the focus of most of the applications
  - Combined with optimization techniques it can be useful in many situations
  - Two-player is most common
- There is a big focus on static bargaining models
  - Very few applications looked at dynamic bargaining
- Controversy on the fairness of the NBS in wireless networks especially
  - It is really fair or do we use just because we “know” it well?

Bargaining Theory in Wireless Networks – Future Directions

- Dynamic bargaining is largely left unexplored in wireless networks
  - Advanced algorithms for dynamic bargaining are of interest in many situations, e.g., spectrum access in cognitive network
- More analysis on the practicality of the bargaining models (static or dynamic) in wireless applications
  - Are we taking into account costs for bargaining?
  - How to implement these solutions in a decentralized network?
- The use of bargaining in conjunction with other techniques such as non-cooperative games or coalitional game theory
- New applications
  - Physical layer security: any connection between secrecy rates and Nash bargaining or KSBS?
  - Smart grid: allocation of utility resources?
  - Routing in wireless networks

Introduction

- Introduction to cooperative game
- Bargaining solution
  - Nash Bargaining Solution
  - Kalai – Smorodinsky Bargaining Solution
  - Rubinstein Bargaining Process
  - Example of Bargaining in Wireless Networks
- Coalitional game
  - Class I: Canonical coalitional games
  - Class II: Coalition formation games
  - Class III: Coalitional graph games
- Summary

Coalitional Games: Preliminaries

- **Definition** of a coalitional game (N,v)
  - A set of players ***N*****,** **a** ***coalition*** S is a group of cooperating players (subset of N)
  - Worth (utility) of a coalition ***v***
    - *In general,* ***payoff v(S)*** *is a real number that represents the gain resulting from a coalition S in the game (N,v)*
    - ***v(N)*** *is the worth of forming the coalition of all users, known as the* ***grand coalition***
  - User payoff x_i : the portion of ***v(S)*** received by a player ***i*** in coalition S

Coalitional Games: Utility

- Transferable utility (TU)
  - The worth ***v(S)*** of a coalition *S* can be distributed arbitrarily among the players in a coalition hence,
  - ***v(S)*** is a **function** from the power set of N over the real line
- Non-transferable utility (NTU)
  - The payoff that a user receives in a coalition is pre-determined, and hence the value of a coalition cannot be described by a function
  - ***v(S)*** is a set of payoff vectors that the players in *S* can achieve

$$v(S) \subseteq \mathbb{R}^{|S|}$$

  - Developed by Auman and Peleg (1960) using a non-cooperative game in strategic form as a basis

Payoff division

- Equal fair
  - Each user guarantees its non-cooperative utility
  - The extra worth is divided equally among coalition users
- Proportional fair
  - Each user guarantees its non-cooperative utility
  - A proportional fair division, based on the non-cooperative worth, is done on the extra utility available through cooperation
- Other fairness
  - Shapley value
  - Nucleolus
  - Market Fairness

An example coalitional game

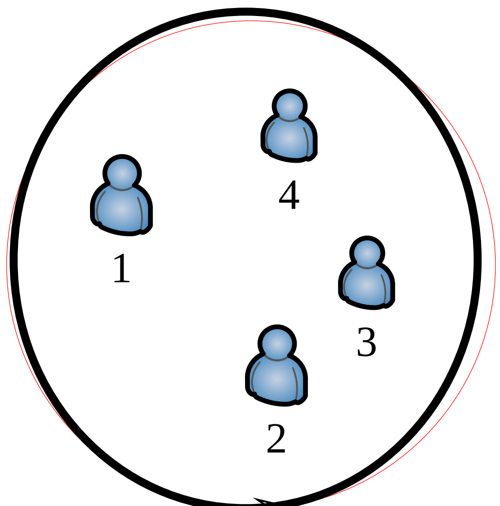
- Example of a coalition game: Majority Vote

$$v(S) = \begin{cases} 1, & \text{if } |S| > |N|/2; \\ 0, & \text{otherwise.} \end{cases}$$

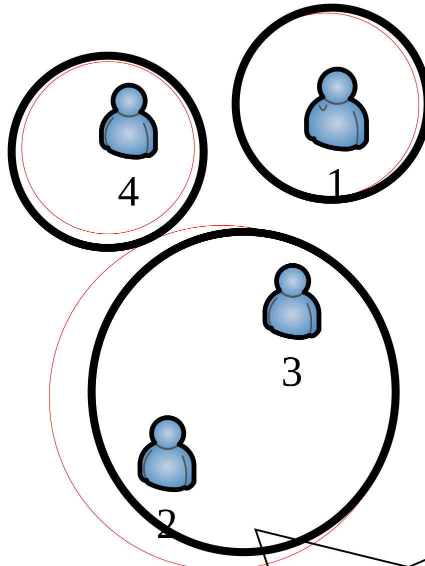
  - President is elected by majority vote
  - A coalition consisting of a majority of players has a worth of 1 since it is a decision maker
  - Value of a coalition does not depend on the external strategies of the users
    - *This game is in characteristic function form*
  - If the voters divide the value as money
    - *Transferable utility*

A new classification

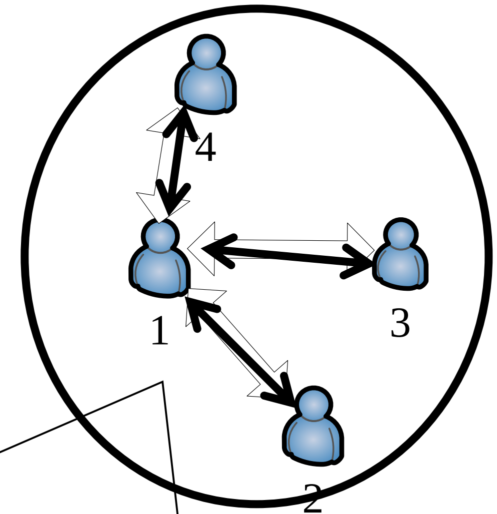
Class I: Canonical Coalitional Games



Class II: Coalition Formation Games



Class III: Coalitional Graph Games



- The **grand** coalition of all users is an optimal structure.
-Key question "How to stabilize the grand coalition?"
- Several well-defined solution concepts exist.

- The network structure that forms depends on gains and costs from cooperation.
-Key question "How to form an appropriate coalitional structure (topology) and how to study its properties?"
- More complex than Class I, with no formal solution concepts.

- Players' interactions are governed by a communication graph structure.
-Key question "How to stabilize the grand coalition or form a network structure taking into account the communication graph?"
- Solutions are complex, combine concepts from coalitions, and non-cooperative games

Class I: Canonical Coalitional Games

- Main properties
  - Cooperation is **always** beneficial
    - *The **grand coalition** is guaranteed to form*
  - The game is **superadditive**

$$v(S_1 \bigcup S_2) \geq v(S_1) + v(S_2) \ \forall S_1, S_2 \subset N \text{ with } S_1 \bigcap S_2 = \emptyset$$

  - **The most famous type of coalitional games!**

- Main Objectives
  - Study the properties and **stability** of the grand coalition
    - *How can we stabilize the grand coalition?*
  - How to divide the utility and gains in a fair manner ?
    - *Improper payoff division => incentive for players to leave coalition*

Canonical games: Solution concepts

- The Core: the most **renowned** concept
  - For a TU game, the core is a set of payoff allocation ***($x_1, \ldots, x_N$)*** satisfying two conditions

    1. $\sum_{i \in N} x_i = v(N)$

    2. $\sum_{i \in S} x_i \geq v(S), \forall S \in N$

  - The core can be empty
    - ***A non-empty core in a superadditive game => stable grand coalition***

- The drawbacks of the core
  - The core is often empty.
  - When the core is non-empty it is often a large set.
  - The allocations that lie in the core are often unfair.

Finding the Core (1)

- **Example 1:** A certain painting is worth a_i dollars to any player *i* in a 3-player coalitional game. Assume $a_1 < a_2 < a_3$, so Player 3 values the object the most. However, the painting is owned by Player 1 and thus $v(\{1\}) = a_1$ while $v(\{2\}) = v(\{3\}) = 0$.
  - Write down the values for the grand coalition and coalitions{2,3}, {1,3}, {1,2}
    - $v(\{2,3\}) = 0$
    - $v(\{1,2\}) = a_2$
    - $v(\{1,3\}) = a_3 = v(\{1,2,3\})$
  - Now let's try to find the core

Finding the Core (2)

- The core of **Example 1**

$$
\begin{array}{lll}
x_1 \geq a_1 & x_1 + x_2 \geq a_2 & \\
x_2 \geq 0 & x_1 + x_3 \geq a_3 & x_1 + x_2 + x_3 = a_3 \\
x_3 \geq 0 & x_2 + x_3 \geq 0 &
\end{array}
$$

- Through these inequalities we can deduce that $x_2 = 0$, $x_1 \geq a_2$ *and* $x_3 = a_3 - x_1$
- An allocation in the core would, thus, be in the form $\mathbf{(x, 0, a_3 - x)}$ **with** $\mathbf{a_2 \leq x \leq a_3}$
  - Player 3 buys the painting from Player 1 at a price x
  - Player 2 does not get any payoff but its participation forced the price to be more than a_2

Finding the Core (3)

- Graphical Method
  - Provides an insight on the core existence
  - Suited for small games (typically 3-player games)
- **Example 2**
  - Consider a 3 players game with N={1,2,3}

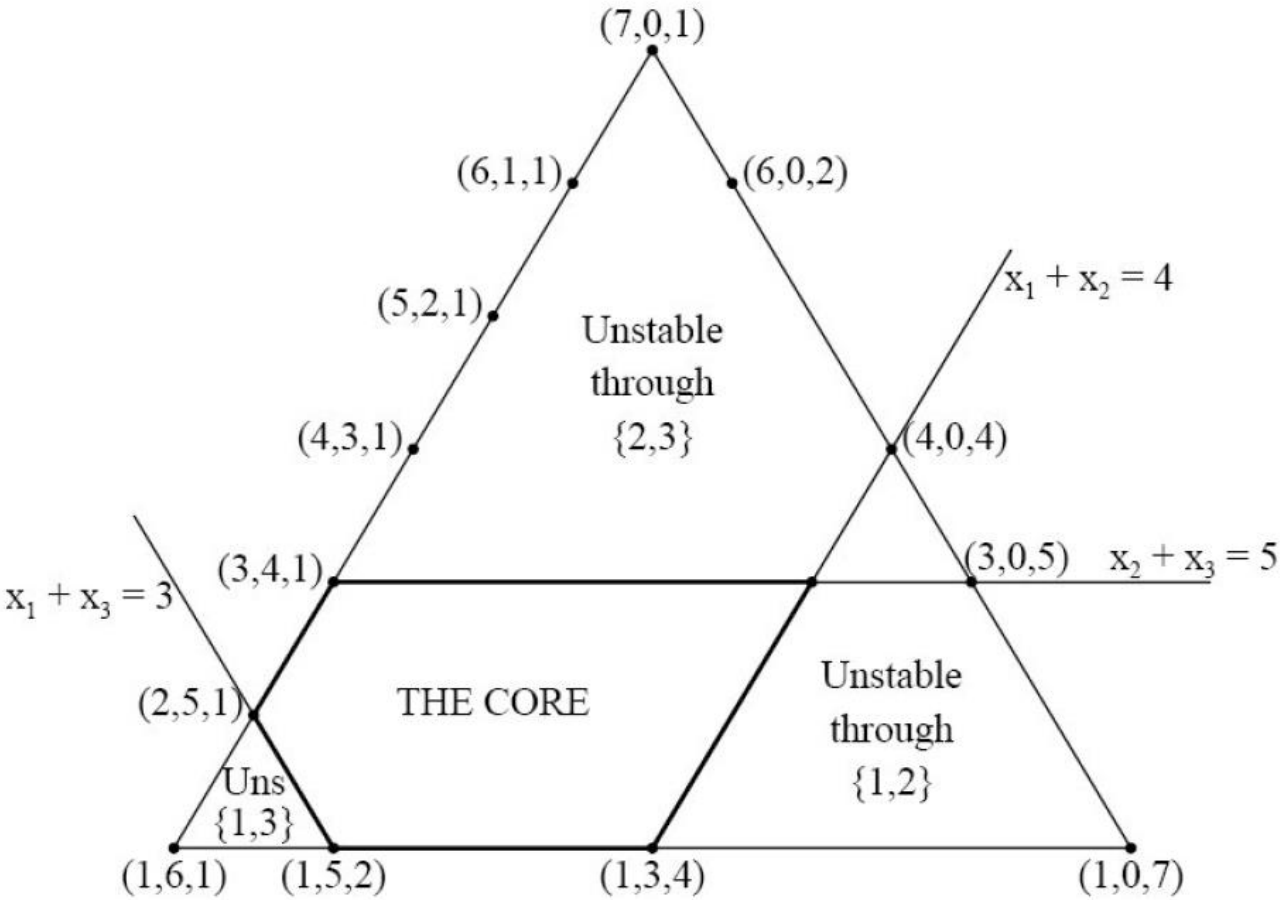
$$v(\{1\}) = v(\{3\}) = 1 \qquad v(\{2\}) = 0$$

$$v(\{1,2\}) = 4 \quad v(\{1,3\}) = 3 \quad v(\{2,3\}) = 5$$

$$v(\{1,2,3\}) = 8 \qquad .$$

Finding the Core (4)

- Plotting the constraints in the plane $x_1 + x_2 + x_3 = 8$



Finding the Core (5)

- A TU coalitional game is balanced if

$$\sum_{S \subseteq \mathcal{N}} \mu(S) v(S) \leq v(\mathcal{N})$$

For every balanced collection of weights

$$\sum_{S \supseteq i} \mu(S) = 1, \ \forall i \in \mathcal{N}$$

- Every player has a unit of time that he can distribute among its possible coalitions.
- A game is **balanced** if there is no feasible allocation of time which can yield more than the grand coalition's worth v(*N*)
- Can be hard to prove in certain wireless applications

Finding the Core (6)

- The problem of finding the core can be cast into an LP

$$\min_{\boldsymbol{x}} \sum_{i \in \mathcal{N}} x_i$$

$$\text{s.t.} \sum_{i \in S} x_i \geq v(S), \ \forall S \subseteq \mathcal{N}.$$

- Using some duality results, we have the following theorem
  - (Bondareva-Shapley) The core of a game is non-empty, if and only if the game is **balanced**

Finding the Core (7)

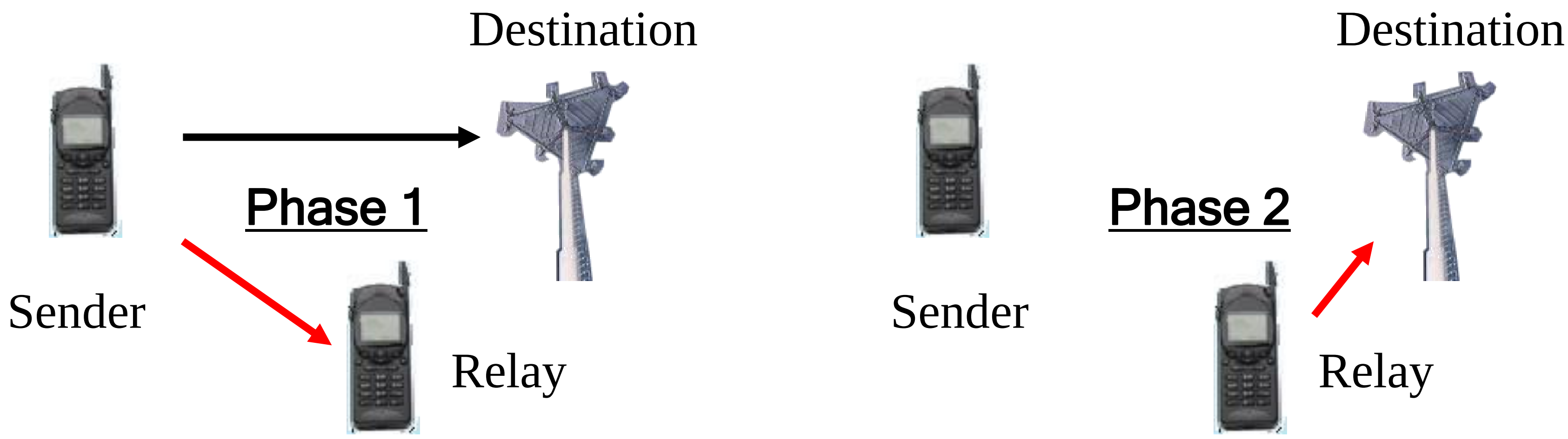
- A TU coalitional game is **convex** if

$$v(S_1) + v(S_2) \leq v(S_1 \cup S_2) + v(S_1 \cap S_2) \; \forall \; S, T \subseteq \mathcal{N}.$$

- A convex game has a non-empty core
  - Convexity is hard to have but has some interesting results (later slides)
- In simple games (i.e. games where the value is either 0 or 1), if a veto player exists, then the core is non-empty
- Other ways to prove the core existence
  - Checking whether fair allocation lie in the core
  - Using application-specific techniques (information theory, etc)
    - *Some insights in next lecture*

Ex: Cooperative Transmission

- New communication paradigm
  - Exploring broadcast nature of wireless channel
  - Relays can be served as virtual antenna of the source
  - MIMO system
  - Multi-user and multi-route diversity



  - Most popular research in current wireless communication
  - Industrial standard: IEEE WiMAX 802.16J

Cooperative Transmission Model

- No cooperation (direct transmission), primary user needs power P_d.
- Cooperative transmission
  - Stage one: direct transmission. s, source; r, relay; d, destination

$$y_{s,d} = \sqrt{P_0} h_{s,d} x + n_{s,d},$$

and $y_{s,r_i} = \sqrt{P_0} h_{s,r_i} x + n_{s,r_i}, \forall i \in \{1, \dots, N\}$

  - Stage two: relay retransmission using orthogonal channels, amplified-and-forward

$$y_{r_i,d} = \frac{\sqrt{P_i}}{\sqrt{P_0 |h_{s,r_i}|^2 + \sigma^2}} h_{r_i,d} y_{s,r_i} + n_{r_i,d}.$$

  - Maximal ration combining at the receiver of backbone node

$$\Gamma = \Gamma_0 + \sum_{i=1}^{N} \Gamma_i \quad \Gamma_0 = \frac{P_0 |h_{s,d}|^2}{\sigma^2} \quad \Gamma_i = \frac{P_0 P_i |h_{s,r_i}|^2 |h_{r_i,d}|^2}{\sigma^2 (P_0 |h_{s,r_i}|^2 + P_i |h_{r_i,d}|^2 + \sigma^2)}$$

  - To achieve same SNR, power saving for primary user $P_0 < P_d$

Main Idea

To get a good position, try to volunteer first

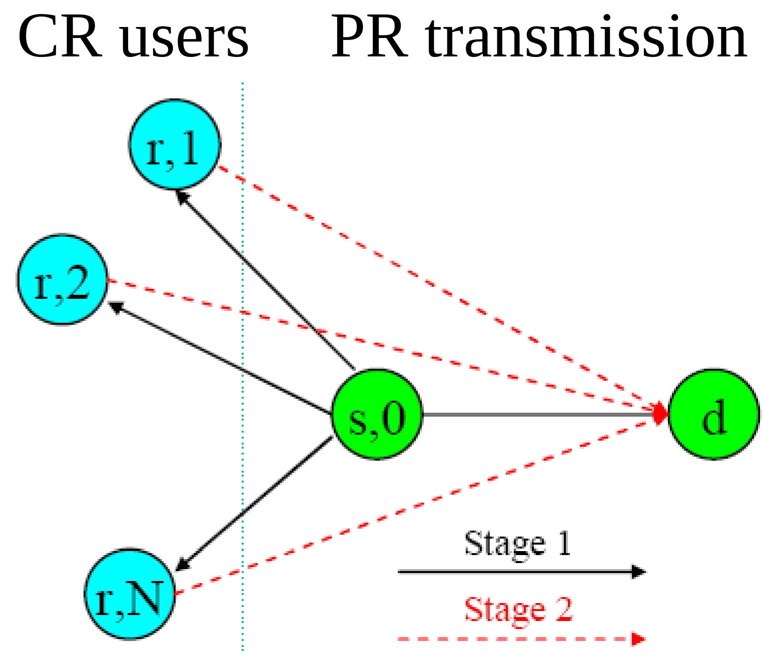


Fig. 2: Coalition Game with Cooperative Transmission

- CR nodes help the PU node reduce transmission power using cooperative transmission, for future rewards of transmission.
- The idea can be formulated by a coalition game.

Other applications of canonical games

- Zhu Han and H. Vincent Poor, ``Coalition Games with Cooperative Transmission: A Cure for the Curse of Boundary Nodes in Selfish Packet-Forwarding Wireless Networks", IEEE Transactions on Communications. vol. 57, No. 1, P.P. 203-213, January 2009.
- Rate allocation in a Gaussian multiple access channel (La and Anantharam, 2003)
  - The grand coalition maximizes the channel capacity
  - How to allocate the capacity in a fair way that stabilizes the grand coalition?
    - *The Core, Envy-free fairness (a variation on the Shapley value)*
- Vitual MIMO (W. Saad, Z. Han, M. Debbah, A. Hjorungnes, 2008)
- Allocation of channels in a cognitive radio network when service providers cooperate in a grand coalition (Aram et al., INFOCOM, 2009)
- Any application where
  - The grand coalition forms (no cost for cooperation)
  - Stability and fairness are key issues

Class II: Coalition Formation Games

- Main Properties
  - The game is NOT **superadditive**
  - Cooperation gains are limited by a cost
    - *The **grand coalition** is NOT guaranteed to form*
  - Cluster the network into partitions
  - **New issues:** network topology, coalition formation process, environmental changes, etc

- Key Questions
  - How can the users form coalitions?
  - What is the network structure that will form?
  - How can the users adapt to environmental changes such as mobility, the deployment of new users, or others?
  - Can we say anything on the stability of the network structure?

Coalition Formation: Merge and Split

- **Merge rule**: merge any group of coalitions where

$$\{\cup_{j=1}^{l} S_j\} \triangleright \{S_1, \ldots, S_l\}$$

- **Split rule**: split any group of coalitions where

$$\{S_1, \ldots, S_l\} \triangleright \{\cup_{j=1}^{l} S_j\}$$

- A decision to merge (split) is an agreement between all players to form (break) a new coalition
  - Socialist (social well fare improved by the decision)
  - Capitalist (individual benefit improved)

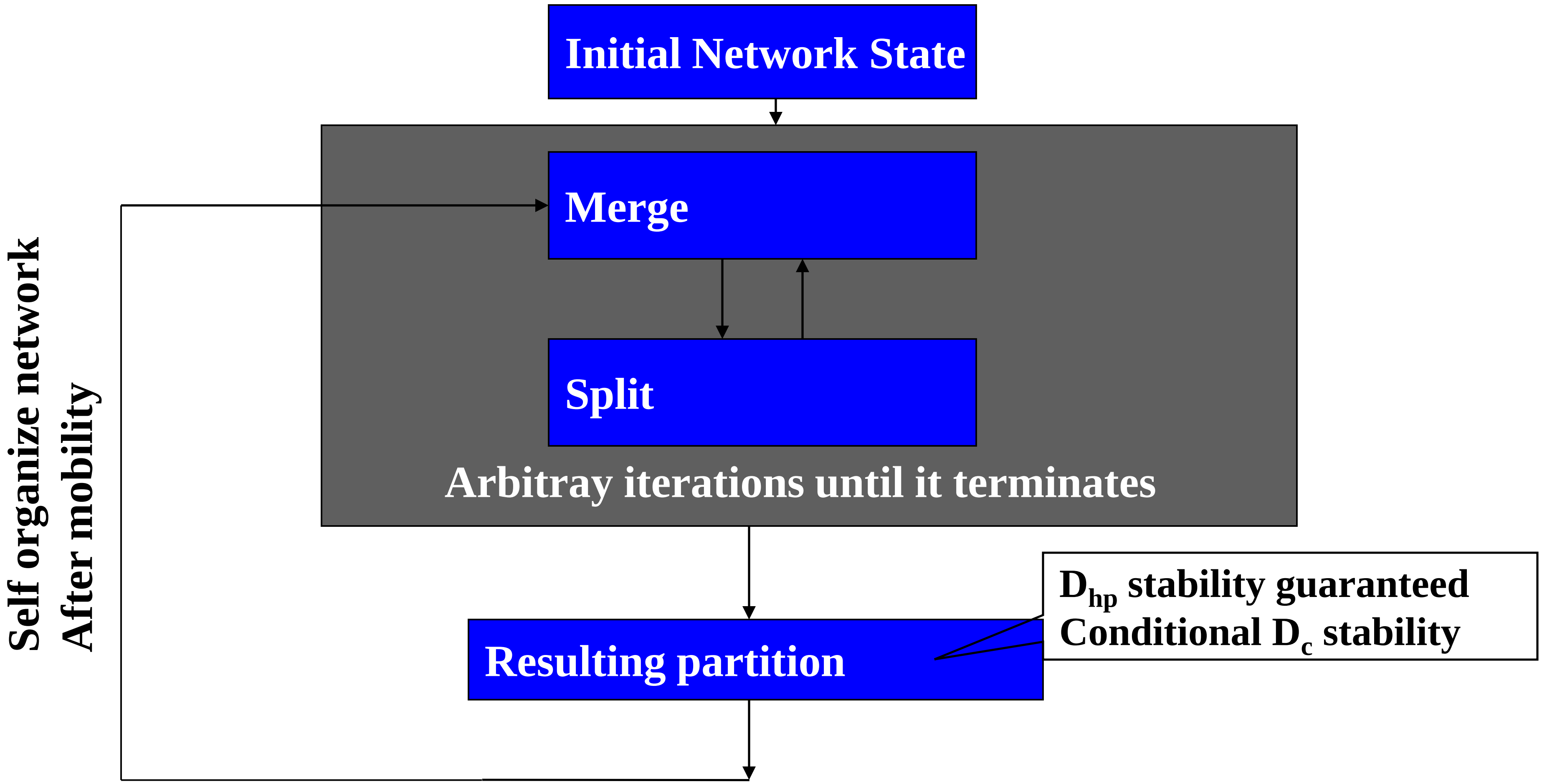
Merge and Split: Properties

- Any merge and split iteration converges and results in a final partition.
- Merge and split decision
  - Individual decision
  - Coalition decision
  - Can be implemented in a distributed manner with no reliance on any centralized entity
- Using the Pareto order ensures that no player is worse off through merge or split
  - Other orders or preference relations can be used

Stability Notions

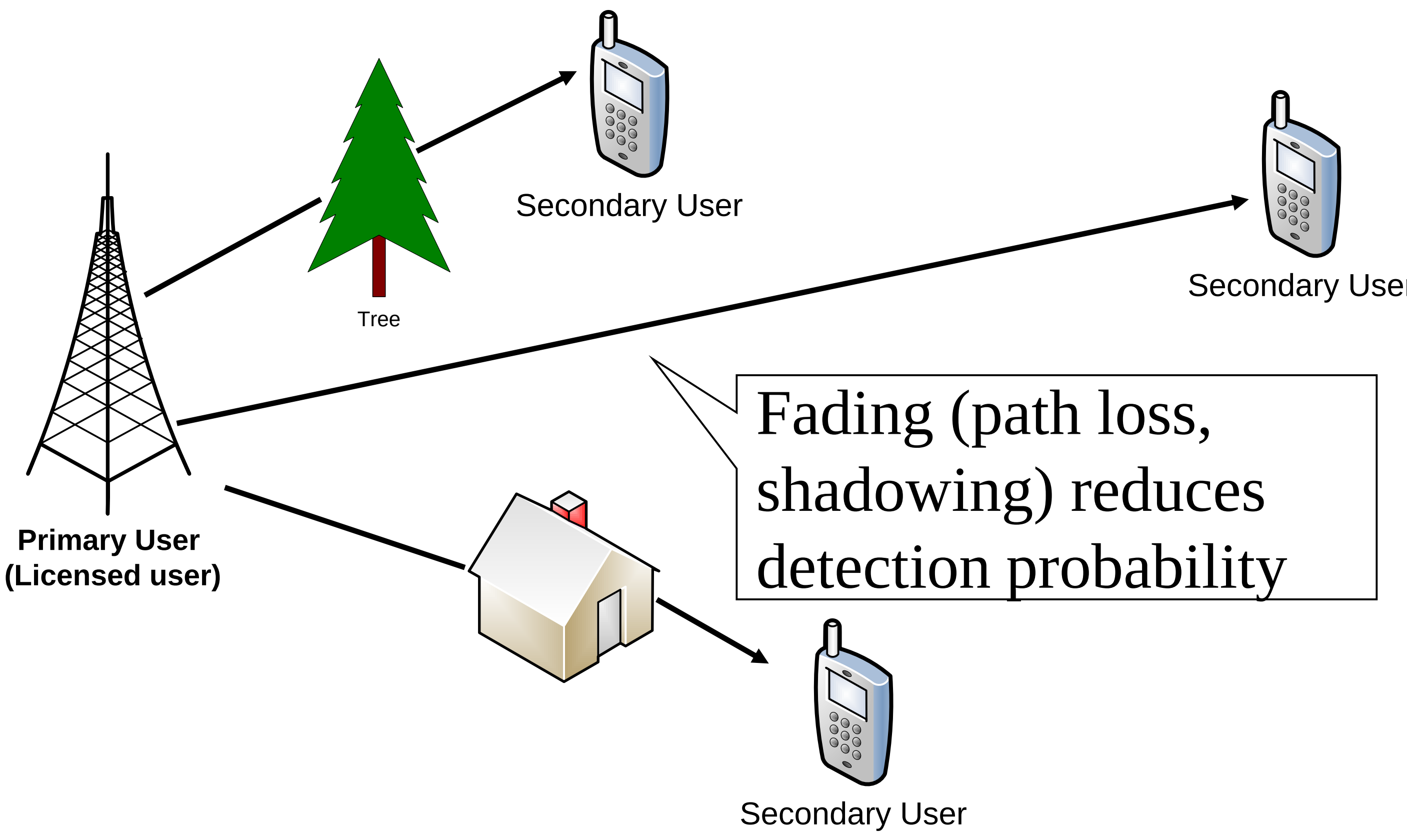
- D_{hp} stable
  - No users can defect via merge/split
  - Partition resulting from merge and split is D_{hp} stable
- D_c stable
  - No users can defect to form a new collection in N
  - A D_c stable partition is **socially optimal**
  - When it exists, it is the **unique** outcome of any merge and split iteration
  - Strongest type of stability

Merge and Split algorithm

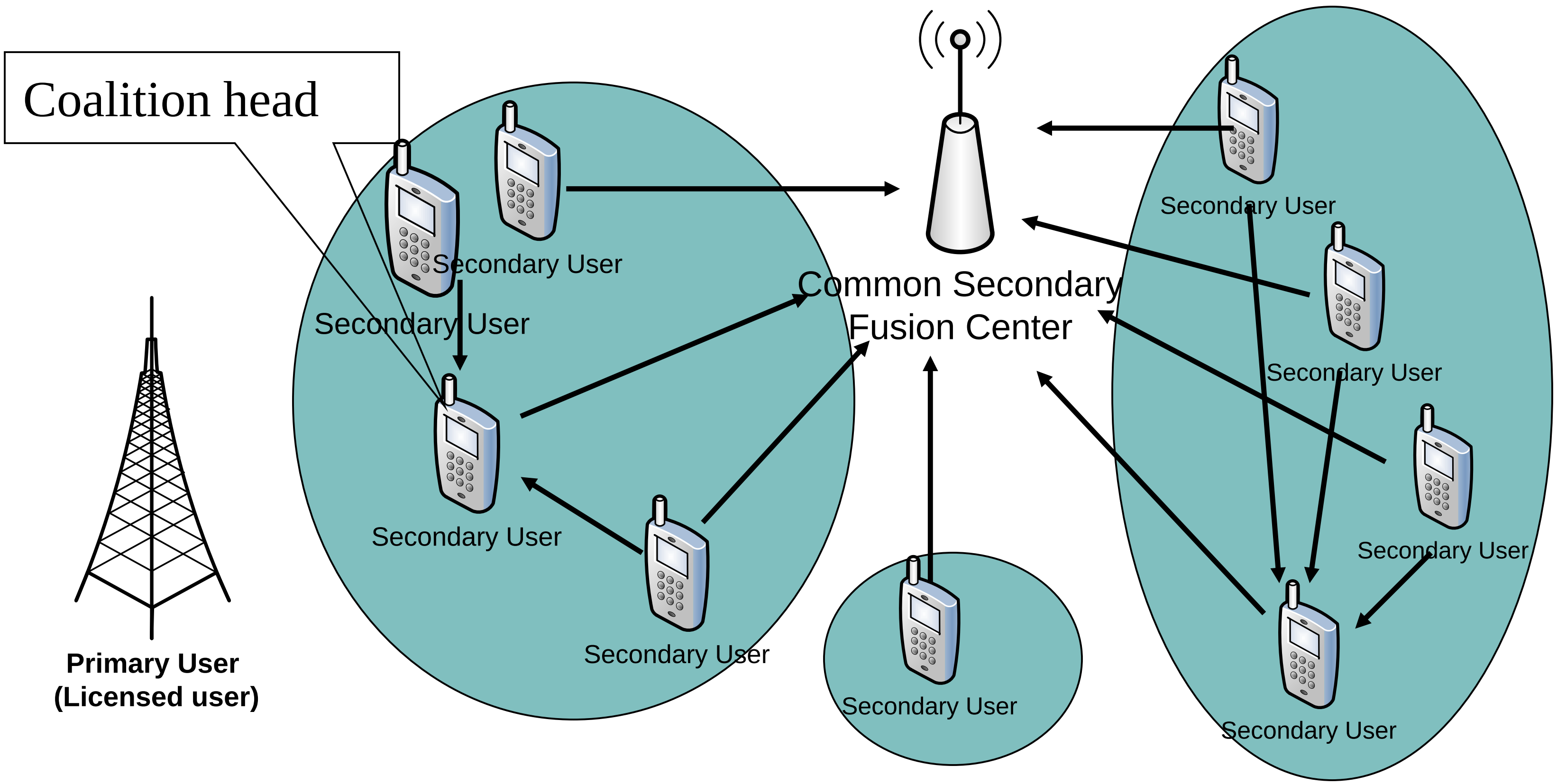


Hidden Terminal Problem

- In fading environments, local sensing suffers from hidden terminal problem



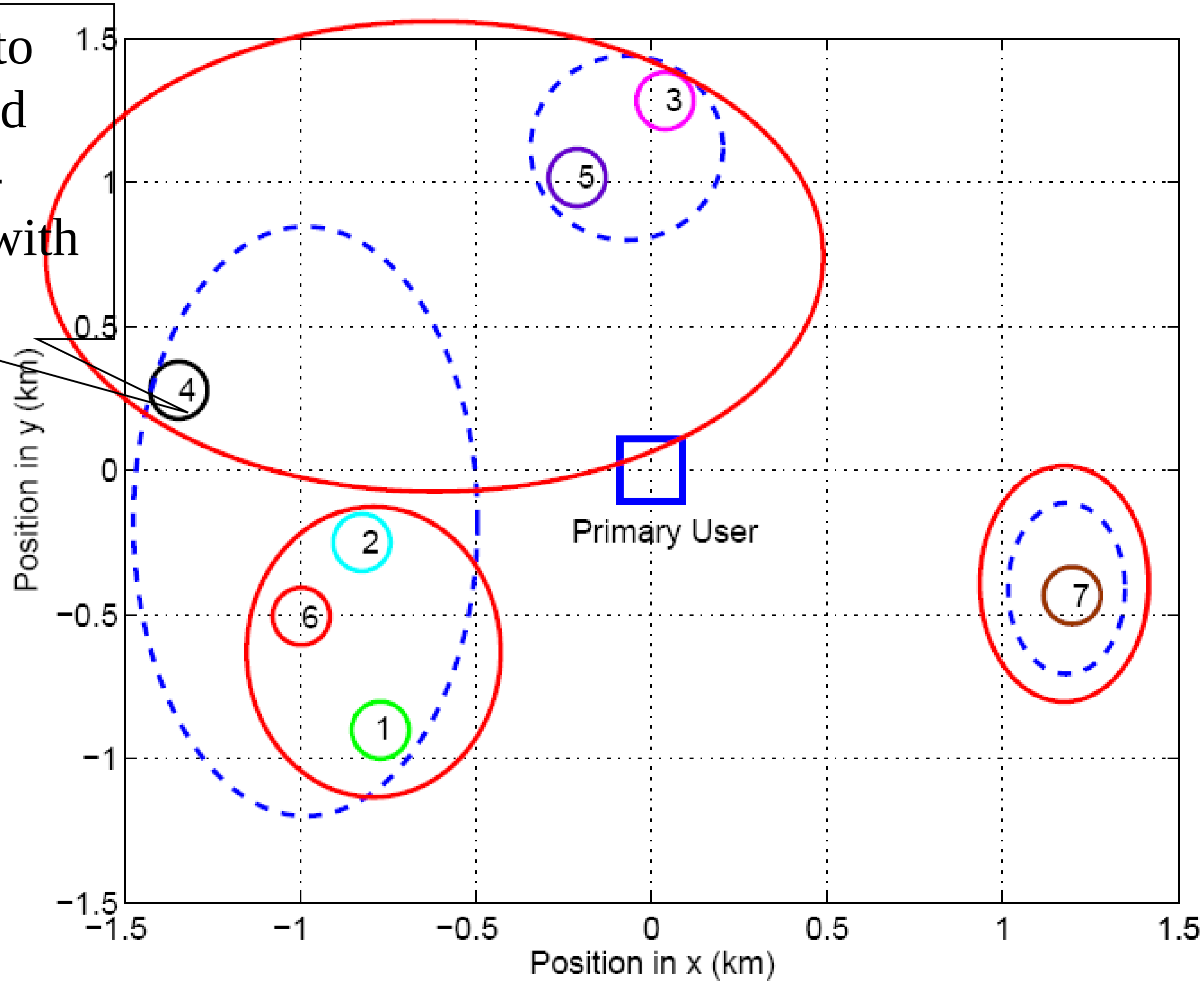
Distributed Collaborative Sensing



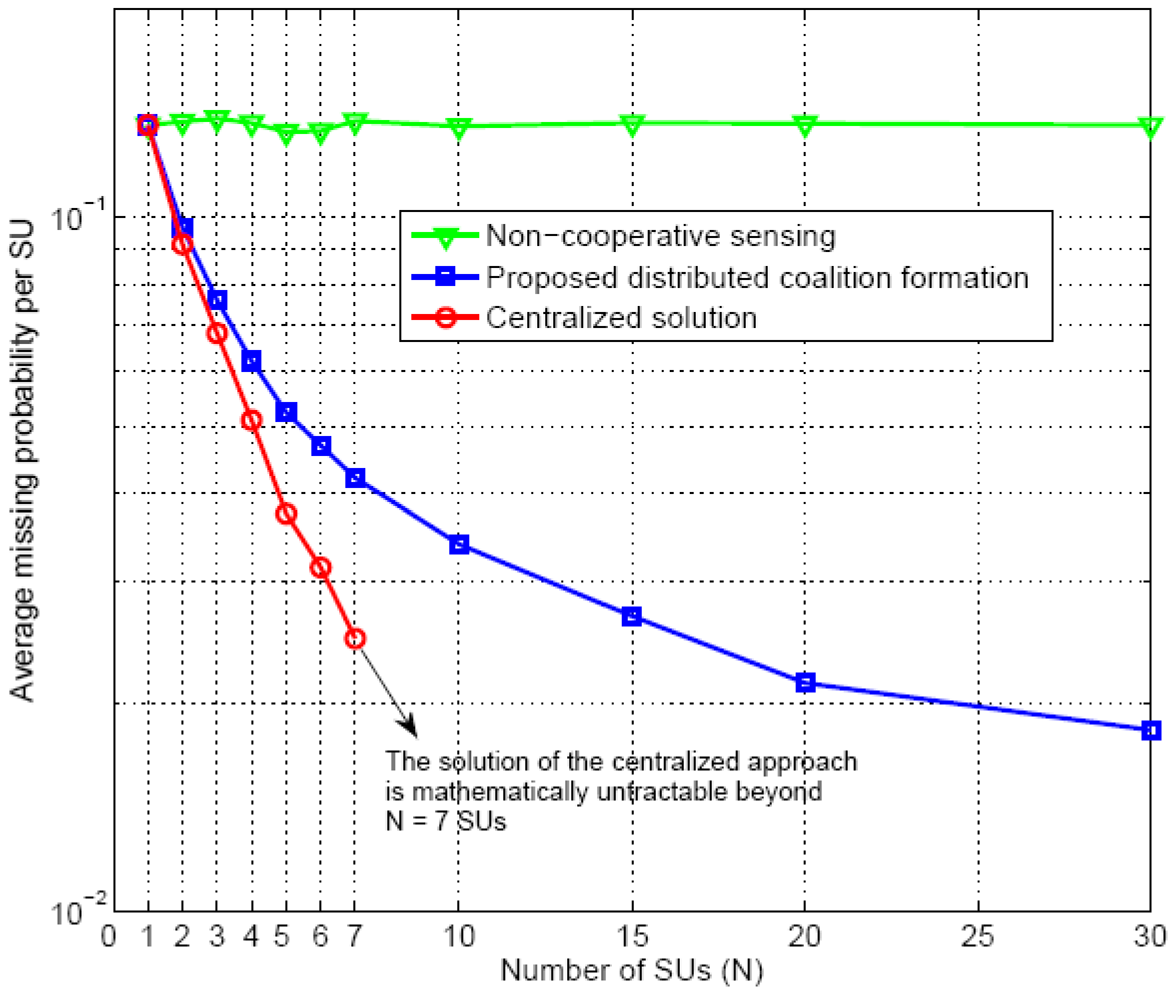
- Distributed collaborative sensing between the users with **no centralized** fusion center
- Which groups will form?
  - **Coalitional games!**

Simulation Results

When allowed to make distributed decisions, SU 4 prefers to stay with {2,1,6}

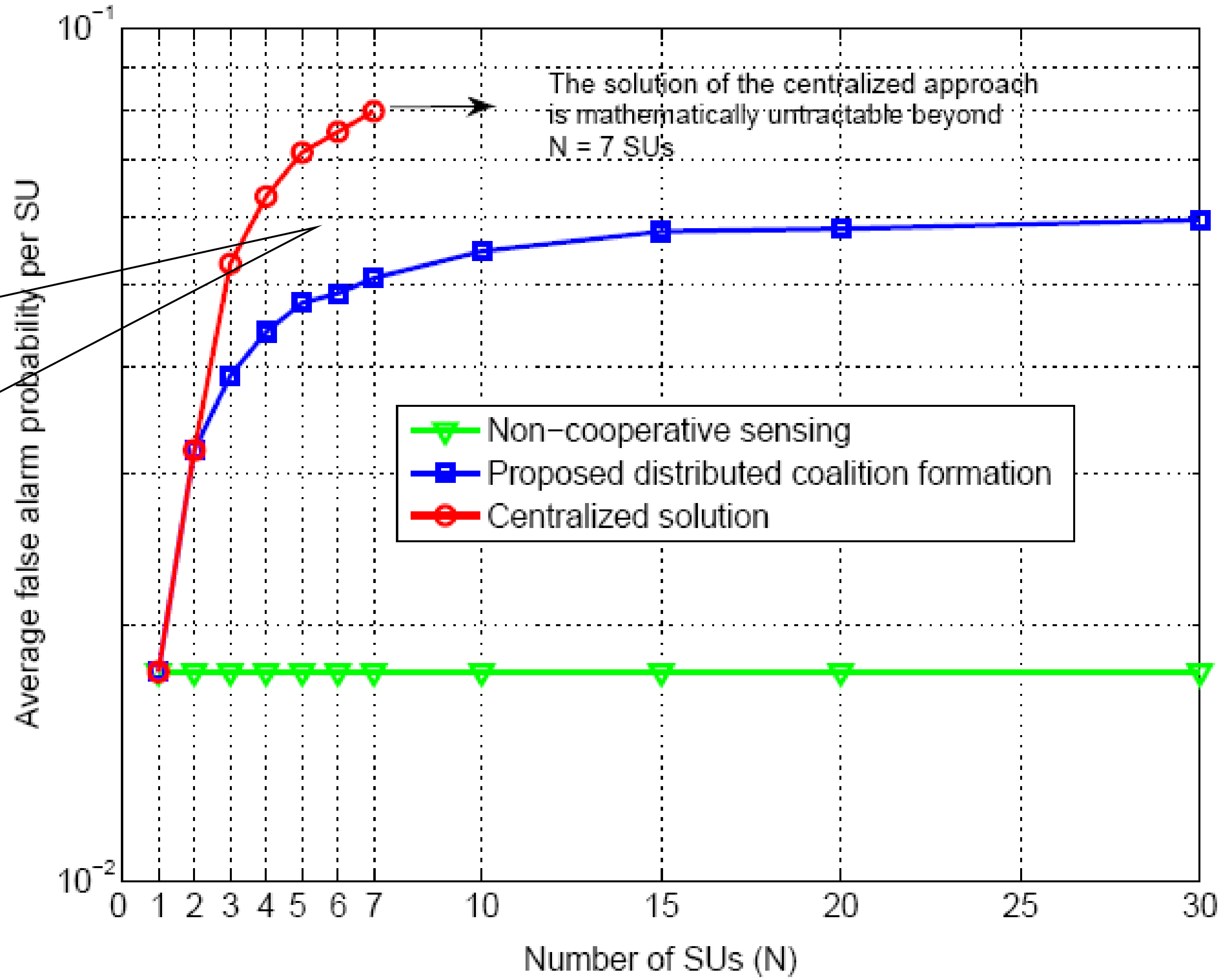


Simulation Results (1)



Simulation Results (2)

The gap with the optimal solution in probability of miss performance is compensated by a lower false alarm

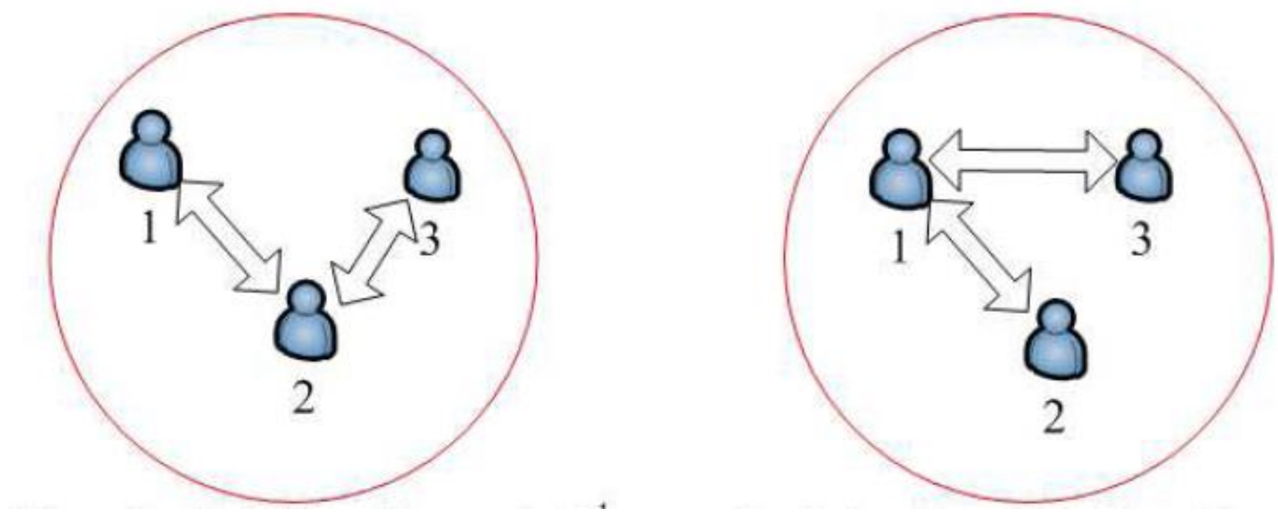


Other applications of coalition formation

- Coalitional games for topology design in wireless networks
  - **Physical layer security**
    - *Merge-and-split for improving secrecy capacity*
    - ***W. Saad, Z. Han, T. Basar, M. Debbah and A. Hjørungnes, "Physical layer security: coalitional games for distributed cooperation," WiOpt, 2009***
  - **Task allocation among UAVs in wireless networks**
    - *Hedonic coalition formation*
    - ***W. Saad, Z. Han, T. Basar, M. Debbah and A. Hjørungnes, "A selfish approach to coalition formation in wireless networks," GameNets, 2009***
  - **Vehicular Network**
    - *``Coalition Formation Games for Distributed Roadside Units Cooperation in Vehicular Networks", JSAC Jan. 2011*
  - **Endless possibilities**
    - *Study of cooperation when there is cooperation with cost*
    - *Topology design in wireless networks*
    - *Beyond wireless: smart grid*

Class III: Coalition Graph Games

- Main properties
  - The game is in **graph form**
    - *May depend on externalities also*
  - There is a **graph** that connects the players of every coalition
  - Cooperation with or without cost
  - **A Hybrid type of games:** concepts from classes I and II, as well as non-cooperative games



Coalition Graph Games

- First thought of by Myerson, 1977, called "Coalitional games with communication structure"
  - Axiomatic approach to find a Shapley-like value for a coalitional game with an underlying graph structure
  - Coalition value depends on the graph
  - The dependence is only based on **connections**
- Key Questions
  - How can the users form the **graph** structure that will result in the network?
  - If all players form a single graph (grand coalition with a graph), can it be stabilized?
  - How can the users adapt to environmental changes such as mobility, the deployment of new users, or others?
  - What is the effect of the graph on the game?

Applications of Coalitional Graph Games

- Coalitional graph games for network formation
  - **WiMAX IEEE 802.16j/LTE**
    - *Network formation game for uplink tree structure formation*
    - ***W. Saad, Z. Han, M. Debbah, and A. Hjørungnes, "Network formation games for distributed uplink tree construction in IEEE 802.16j," in proc. GLOBECOM 2008***
    - ***W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "A game-based self-organizing uplink tree for VoIP services in IEEE 802.16j," ICC 2009***
  - **Routing in communication networks**
    - *See the work by Johari (Stanford)*
  - **Many future possibilities**
    - *The formation of graphs is ubiquitous in the context of communication networks*

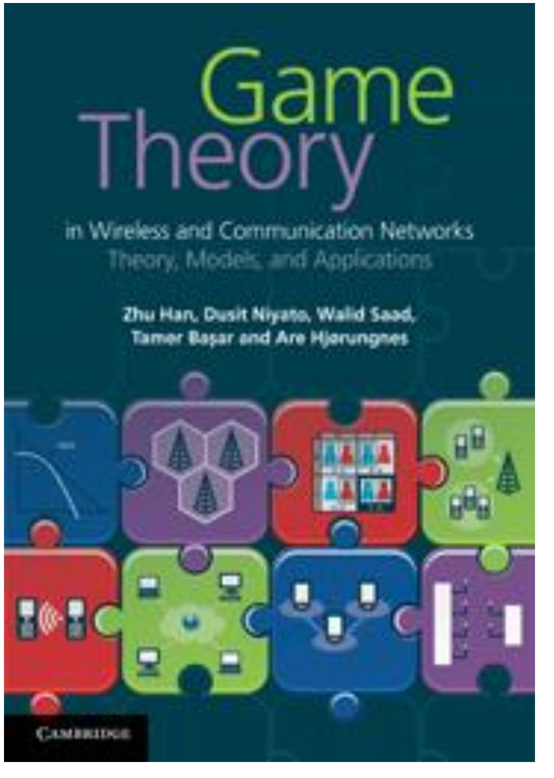
Summary of coalitional game

- Coalitional games are a strong tool for different models in wireless and communication networks
- Bagaining solution for local improvement
- A novel classification that can help in identifying potential applications
- A tool for next generation self-organizing networks
  - Especially through coalition formation and network formation games

Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 3
Differential Game

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of Lecture Notes

- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, book 2
- Contract Theory, Lecture 8, book 2
- Stochastic Game, Lecture 9, book 2
- Learning in Game, Lecture 10, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 11, book 2
- Mean Field Game, Lecture 12, book 2
- Zero Determinant Strategy, Lecture 13, book 2
- Network Economy, Lecture 14, book 2
- Game in Machine Learning, Lecture 15, book 2

Introduction

- Basics
- Controllability
- Linear ODE: Bang-bang control
- Linear time optimal control
- Pontryagin's maximum principle
- Dynamic programming
- Dynamic game
- Note: Some parts are not from the book. See some dynamic control book and Basar's dynamic game book for more references.

Basic Problem

- **ODE:** x: state, f: a function, α: control

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\alpha}(t)) & (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}$$

- **Payoff:** r: running payoff, g: terminal payoff

$$P[\boldsymbol{\alpha}(\cdot)] := \int_0^T r(\mathbf{x}(t), \boldsymbol{\alpha}(t))\, dt + g(\mathbf{x}(T))$$

THE BASIC PROBLEM. Our aim is to find a control $\boldsymbol{\alpha}^*(\cdot)$, which *maximizes* the payoff. In other words, we want

$$P[\boldsymbol{\alpha}^*(\cdot)] \geq P[\boldsymbol{\alpha}(\cdot)]$$

for all controls $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$. Such a control $\boldsymbol{\alpha}^*(\cdot)$ is called *optimal*.

This task presents us with these mathematical issues:

(i) Does an optimal control exist?

(ii) How can we characterize an optimal control mathematically?

(iii) How can we construct an optimal control?

Example

- Moon lander: Newton's law

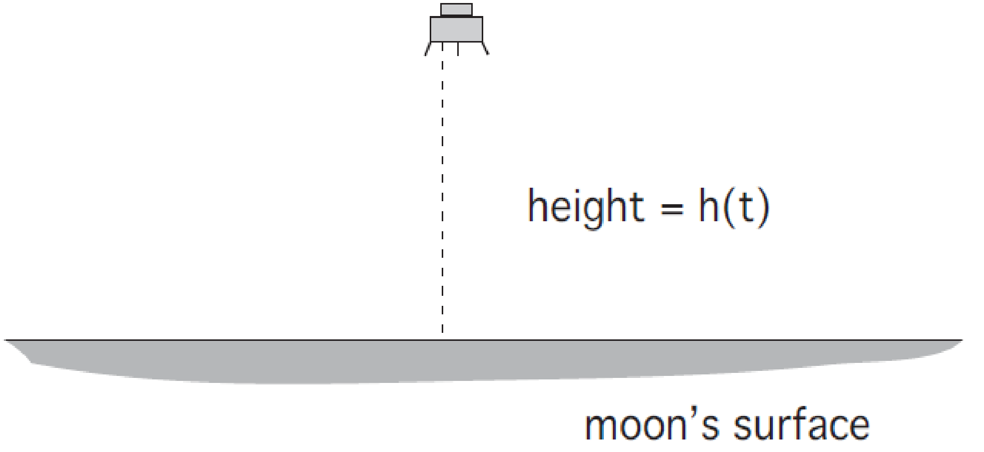
$$\begin{aligned} h(t) &= \text{height at time } t \\ v(t) &= \text{velocity} = \dot{h}(t) \\ m(t) &= \text{mass of spacecraft (changing as fuel is burned)} \\ \alpha(t) &= \text{thrust at time } t \end{aligned}$$

$$m\ddot{h} = -gm + \alpha$$

- ODE

$$\begin{cases} \dot{v}(t) = -g + \frac{\alpha(t)}{m(t)} \\ \dot{h}(t) = v(t) \\ \dot{m}(t) = -k\alpha(t). \end{cases}$$

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \alpha(t)) \\ \mathbf{x}(t) &= (v(t), h(t), m(t)) \end{aligned}$$



- Objective: minimize fuel

Maximize the remain

$$P[\alpha(\cdot)] = m(\tau)$$

- Constraints $h(t) \geq 0, \quad m(t) \geq 0$

Controllability

CONTROLLABILITY QUESTION: Given the initial point x^0 and a "target" set $S \subset \mathbb{R}^n$, does there exist a control steering the system to S in finite time?

DEFINITION. We define the *reachable set for time t* to be

$$\mathcal{C}(t) = \text{set of initial points } x^0 \text{ for which there exists a control such that } \mathbf{x}(t) = 0,$$

and the overall *reachable set*

$$\mathcal{C} = \text{set of initial points } x^0 \text{ for which there exists a control such that } \mathbf{x}(t) = 0 \text{ for some finite time } t.$$

Note that

$$\mathcal{C} = \bigcup_{t \geq 0} \mathcal{C}(t).$$

Linear ODE

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t) & (t > 0) \\ \mathbf{x}(0) = x^0, \end{cases}$$

where $M \in \mathbb{M}^{n\times n}$ and $N \in \mathbb{M}^{n\times m}$. We assume the set A of control parameters is a cube in $\mathbb{R}^m$:

$$A = [-1,1]^m = \{a \in \mathbb{R}^m \mid |a_i| \leq 1,\ i = 1,\dots,m\}.$$

THEOREM 2.1 (SOLVING LINEAR SYSTEMS OF ODE).
(i) *The unique solution of the homogeneous system of ODE*

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{x}(0) = x^0 \end{cases}$$

is

$$\mathbf{x}(t) = \mathbf{X}(t)x^0 = e^{tM}x^0.$$

(ii) *The unique solution of the nonhomogeneous system*

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) + \mathbf{f}(t) \\ \mathbf{x}(0) = x^0. \end{cases}$$

is

$$\mathbf{x}(t) = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)\mathbf{f}(s)\,ds.$$

This expression is the *variation of parameters formula.*

Controllability of Linear Equations

DEFINITION. The *controllability matrix* is

$$G = G(M,N) := \underbrace{[N, MN, M^2N, \ldots, M^{n-1}N]}_{n\times(mn)\text{ matrix}}.$$

DEFINITION. We say the linear system (ODE) is *controllable* if $\mathcal{C} = \mathbb{R}^n$.

THEOREM 2.5 (CRITERION FOR CONTROLLABILITY). *Let A be the cube $[-1,1]^m$ in $\mathbb{R}^n$. Suppose as well that* $\operatorname{rank} G = n$, *and* $\operatorname{Re}\lambda < 0$ *for each eigenvalue λ of the matrix M.*

Then the system (ODE) is controllable.

THEOREM 2.6 (IMPROVED CRITERION FOR CONTROLLABILITY). *Assume* $\operatorname{rank} G = n$ *and* $\operatorname{Re}\lambda \leq 0$ *for each eigenvalue λ of M.*

Then the system (ODE) is controllable.

Observability

- Observation $\mathbf{y}(t) := N\mathbf{x}(t) \qquad (t \geq 0)$ for a given matrix $N \in \mathbb{M}^{m\times n}$

DEFINITION. The pair (ODE),(O) is called *observable* if the knowledge of $\mathbf{y}(\cdot)$ on any time interval $[0, t]$ allows us to compute x^0.

THEOREM 2.7 (OBSERVABILITY AND CONTROLLABILITY). *The system*

$$\begin{cases} \dot{\mathbf{x}}(t) = M\mathbf{x}(t) \\ \mathbf{y}(t) = N\mathbf{x}(t) \end{cases} \tag{2.11}$$

is observable if and only if the system

$$\dot{\mathbf{z}}(t) = M^T\mathbf{z}(t) + N^T\alpha(t), \quad A = \mathbb{R}^m \tag{2.12}$$

is controllable, meaning that $\mathcal{C} = \mathbb{R}^n$.

INTERPRETATION. This theorem asserts that somehow "*observability and controllability are dual concepts*" for linear systems.

Bang-bang Control

DEFINITION. A control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$ is called *bang-bang* if for each time $t \geq 0$ and each index $i = 1, \ldots, m$, we have $|\alpha^i(t)| = 1$, where

$$\boldsymbol{\alpha}(t) = \begin{pmatrix} \alpha^1(t) \\ \vdots \\ \alpha^m(t) \end{pmatrix}.$$

THEOREM 2.8 (BANG-BANG PRINCIPLE). *Let $t > 0$ and suppose $x^0 \in \mathcal{C}(t)$, for the system*

$$\dot{\mathbf{x}}(t) = M\mathbf{x}(t) + N\boldsymbol{\alpha}(t).$$

Then there exists a bang-bang control $\boldsymbol{\alpha}(\cdot)$ which steers x^0 to 0 at time t.

- Bang-bang control is optimal

Existence Of Time-optimal Controls

- Minimize the time from any point to the origin

$$\text{(P)} \qquad P[\boldsymbol{\alpha}(\cdot)] := -\int_0^{\tau} 1 \ ds = -\tau,$$

where $\tau = \tau(\boldsymbol{\alpha}(\cdot))$ denotes the first time the solution of our ODE (3.1) hits the origin 0. (If the trajectory never hits 0, we set $\tau = \infty$.)

OPTIMAL TIME PROBLEM: We are given the starting point $x^0 \in \mathbb{R}^n$, and want to find an optimal control $\boldsymbol{\alpha}^*(\cdot)$ such that

$$P[\boldsymbol{\alpha}^*(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot)\in\mathcal{A}} P[\boldsymbol{\alpha}(\cdot)].$$

Then

$$\tau^* = -\mathcal{P}[\boldsymbol{\alpha}^*(\cdot)] \ \text{ is the minimum time to steer to the origin.}$$

THEOREM 3.1 (EXISTENCE OF TIME OPTIMAL CONTROL). *Let $x^0 \in \mathbb{R}^n$. Then there exists an optimal bang-bang control $\boldsymbol{\alpha}^*(\cdot)$.*

Maximum Principle For Linear System

DEFINITION. We define $K(t, x^0)$ to be the *reachable set* for time t. That is,

$$K(t, x^0) = \{x^1 \mid \text{ there exists } \boldsymbol{\alpha}(\cdot) \in \mathcal{A} \text{ which steers from } x^0 \text{ to } x^1 \text{ at time } t\}.$$

Since $\mathbf{x}(\cdot)$ solves (ODE), we have $x^1 \in K(t, x^0)$ if and only if

$$x^1 = \mathbf{X}(t)x^0 + \mathbf{X}(t)\int_0^t \mathbf{X}^{-1}(s)N\boldsymbol{\alpha}(s)\,ds = \mathbf{x}(t)$$

for some control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$.

THEOREM 3.2 (GEOMETRY OF THE SET K). *The set $K(t, x^0)$ is convex and closed.*

THEOREM 3.3 (PONTRYAGIN MAXIMUM PRINCIPLE FOR LINEAR TIME-OPTIMAL CONTROL). *There exists a nonzero vector h such that*

$$\text{(M)} \qquad h^T\mathbf{X}^{-1}(t)N\boldsymbol{\alpha}^*(t) = \max_{a \in A}\{h^T\mathbf{X}^{-1}(t)Na\}$$

for each time $0 \le t \le \tau^*$.

INTERPRETATION. The significance of this assertion is that if we know h then the *maximization principle* (M) provides us with a formula for computing $\boldsymbol{\alpha}^*(\cdot)$, or at least extracting useful information.

Hamiltonian

- Definition $H(x,p,a) := (Mx + Na)\cdot p \qquad (x, p \in \mathbb{R}^n, a \in A)$

THEOREM 3.4 (ANOTHER WAY TO WRITE PONTRYAGIN MAXIMUM PRINCIPLE FOR TIME-OPTIMAL CONTROL). *Let $\boldsymbol{\alpha}^*(\cdot)$ be a time optimal control and $\mathbf{x}^*(\cdot)$ the corresponding response.*

Then there exists a function $\mathbf{p}^*(\cdot) : [0, \tau^*] \to \mathbb{R}^n$, such that

$$(ODE) \qquad \dot{\mathbf{x}}^*(t) = \nabla_p H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)),$$

$$(ADJ) \qquad \dot{\mathbf{p}}^*(t) = -\nabla_x H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)),$$

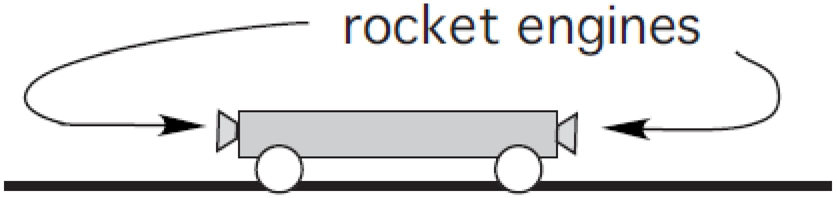
and

$$(M) \qquad H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)) = \max_{a \in A} H(\mathbf{x}^*(t), \mathbf{p}^*(t), a).$$

We call (ADJ) the *adjoint equations* and (M) the *maximization principle*. The function $\mathbf{p}^*(\cdot)$ is the *costate*.

Example: Rocket Railroad Car

- $\mathbf{x}(t) = (q(t), v(t))$ $\quad A = [-1, 1]$



$$\dot{\mathbf{x}}(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{M} \mathbf{x}(t) + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{N} \alpha(t)$$

According to the Pontryagin Maximum Principle, there exists $h \neq 0$ such that

$$(M) \qquad h^T \mathbf{X}^{-1}(t) N \alpha^*(t) = \max_{|a| \leq 1} \{ h^T \mathbf{X}^{-1}(t) N a \}.$$

We will extract the interesting fact that *an optimal control α^* switches at most one time.*

We must compute e^{tM}. To do so, we observe

$$M^0 = I, \; M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; M^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0;$$

and therefore $M^k = 0$ for all $k \geq 2$. Consequently,

$$e^{tM} = I + tM = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Then

$$\mathbf{X}^{-1}(t) = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{X}^{-1}(t) N = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -t \\ 1 \end{pmatrix}$$

$$h^T \mathbf{X}^{-1}(t) N = (h_1, h_2) \begin{pmatrix} -t \\ 1 \end{pmatrix} = -t h_1 + h_2.$$

Example: Rocket Railroad Car

The Maximum Principle asserts

$$(-th_1 + h_2)\alpha^*(t) = \max_{|a|\leq 1}\{(-th_1 + h_2)a\};$$

and this implies that

$$\alpha^*(t) = \operatorname{sgn}(-th_1 + h_2)$$

for the *sign function*

$$\operatorname{sgn} x = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0. \end{cases}$$

Therefore the optimal control α^* switches at most once; and if $h_1 = 0$, then α^* is constant.

Satellite example

Pontryagin Maximum Principle

We come now to the key assertion of this chapter, the theoretically interesting and practically useful theorem that if $\boldsymbol{\alpha}^*(\cdot)$ is an optimal control, then there exists a function $\mathbf{p}^*(\cdot)$, called the *costate*, that satisfies a certain maximization principle. We should think of the function $\mathbf{p}^*(\cdot)$ as a sort of Lagrange multiplier, which appears owing to the constraint that the optimal curve $\mathbf{x}^*(\cdot)$ must satisfy (ODE). And just as conventional Lagrange multipliers are useful for actual calculations, so also will be the costate.

- "The maximum principle was, in fact, the culmination of a long search in the calculus of variations for a comprehensive multiplier rule, which is the correct way to view it: p(t) is a "Lagrange multiplier" . . . It makes optimal control a design tool, whereas the calculus of variations was a way to study nature."

Fixed Time, Free Endpoint Problem

We are given $A \subseteq \mathbb{R}^m$ and also $\mathbf{f} : \mathbb{R}^n \times A \to \mathbb{R}^n$, $x^0 \in \mathbb{R}^n$. We as before denote the set of admissible controls by

$$\mathcal{A} = \{\boldsymbol{\alpha}(\cdot) : [0, \infty) \to A \mid \boldsymbol{\alpha}(\cdot) \text{ is measurable}\}.$$

Then given $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, we solve for the corresponding evolution of our system:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\alpha}(t)) & (t \geq 0) \\ \mathbf{x}(0) = x^0. \end{cases}$$

We also introduce the payoff functional

$$\text{(P)} \qquad P[\boldsymbol{\alpha}(\cdot)] = \int_0^T r(\mathbf{x}(t), \boldsymbol{\alpha}(t))\, dt + g(\mathbf{x}(T)),$$

where the terminal time $T > 0$, running payoff $r : \mathbb{R}^n \times A \to \mathbb{R}$ and terminal payoff $g : \mathbb{R}^n \to \mathbb{R}$ are given.

BASIC PROBLEM: Find a control $\boldsymbol{\alpha}^*(\cdot)$ such that

$$P[\boldsymbol{\alpha}^*(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P[\boldsymbol{\alpha}(\cdot)].$$

Pontryagin Maximum Principle

DEFINITION. The *control theory Hamiltonian* is the function

$$H(x,p,a) := \mathbf{f}(x,a)\cdot p + r(x,a) \qquad (x,p\in\mathbb{R}^n, a\in A).$$

THEOREM 4.3 (PONTRYAGIN MAXIMUM PRINCIPLE). *Assume $\boldsymbol{\alpha}^*(\cdot)$ is optimal for* (ODE), (P) *and $\mathbf{x}^*(\cdot)$ is the corresponding trajectory.*

Then there exists a function $\mathbf{p}^*:[0,T]\to\mathbb{R}^n$ such that

(ODE) $$\dot{\mathbf{x}}^*(t) = \nabla_p H(\mathbf{x}^*(t),\mathbf{p}^*(t),\boldsymbol{\alpha}^*(t)),$$

(ADJ) $$\dot{\mathbf{p}}^*(t) = -\nabla_x H(\mathbf{x}^*(t),\mathbf{p}^*(t),\boldsymbol{\alpha}^*(t)),$$ adjoint equations

and

(M) $$H(\mathbf{x}^*(t),\mathbf{p}^*(t),\boldsymbol{\alpha}^*(t)) = \max_{a\in A} H(\mathbf{x}^*(t),\mathbf{p}^*(t),a) \qquad (0\le t\le \tau^*).$$

In addtion, maximization principle

$$\textit{the mapping} \quad t\mapsto H(\mathbf{x}^*(t),\mathbf{p}^*(t),\boldsymbol{\alpha}^*(t)) \quad \textit{is constant.}$$

Finally, we have the terminal condition

(T) $$\mathbf{p}^*(T) = \nabla g(\mathbf{x}^*(T)).$$ transversality condition

Free Time, Fixed Endpoint Problem

As before, given a control $\boldsymbol{\alpha}(\cdot) \in \mathcal{A}$, we solve for the corresponding evolution of our system:

(ODE)
$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\alpha}(t)) & (t \geq 0) \\ \mathbf{x}(0) = x^0. \end{cases}$$

Assume now that a target point $x^1 \in \mathbb{R}^n$ is given. We introduce then the payoff functional

(P)
$$P[\boldsymbol{\alpha}(\cdot)] = \int_0^\tau r(\mathbf{x}(t), \boldsymbol{\alpha}(t))\, dt$$

Here $r : \mathbb{R}^n \times A \to \mathbb{R}$ is the given running payoff, and $\tau = \tau[\boldsymbol{\alpha}(\cdot)] \leq \infty$ denotes the first time the solution of (ODE) hits the target point x^1.

As before, the basic problem is to find an optimal control $\boldsymbol{\alpha}^*(\cdot)$ such that

$$P[\boldsymbol{\alpha}^*(\cdot)] = \max_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P[\boldsymbol{\alpha}(\cdot)].$$

Pontryagin Maximum Principle

THEOREM 4.4 (PONTRYAGIN MAXIMUM PRINCIPLE). *Assume* $\boldsymbol{\alpha}^*(\cdot)$ *is optimal for* (ODE), (P) *and* $\mathbf{x}^*(\cdot)$ *is the corresponding trajectory.*

Then there exists a function $\mathbf{p}^* : [0, \tau^*] \to \mathbb{R}^n$ *such that*

(ODE) $$\dot{\mathbf{x}}^*(t) = \nabla_p H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)),$$

(ADJ) $$\dot{\mathbf{p}}^*(t) = -\nabla_x H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)),$$

and

(M) $$H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)) = \max_{a \in A} H(\mathbf{x}^*(t), \mathbf{p}^*(t), a) \qquad (0 \le t \le \tau^*).$$

Also,

$$H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t)) \equiv 0 \qquad (0 \le t \le \tau^*).$$

Here τ^* denotes the first first time the trajectory $x^*(\cdot)$ hits the target point x^1. We call $\mathbf{x}^*(\cdot)$ the *state* of the optimally controlled system and $\mathbf{p}^*(\cdot)$ the *costate*.

Example: Linear-quadratic Regulator

(ODE)
$$\begin{cases} \dot{x}(t) = x(t) + \alpha(t) \\ x(0) = x^0, \end{cases}$$

with the quadratic cost functional

$$\int_0^T x(t)^2 + \alpha(t)^2 \, dt,$$

which we want to *minimize*. So we want to maximize the payoff functional

(P)
$$P[\alpha(\cdot)] = -\int_0^T x(t)^2 + \alpha(t)^2 \, dt.$$

For this problem, the values of the controls are not constrained; that is, $A = \mathbb{R}$.

Introducing the Maximum Principle

$$f(x,a) = x + a,\ \ g \equiv 0,\ \ r(x,a) = -x^2 - a^2;$$

and hence

$$H(x,p,a) = fp + r = (x+a)p - (x^2 + a^2)$$

The maximality condition becomes

(M) $$H(x(t), p(t), \alpha(t)) = \max_{a \in \mathbb{R}} \{-(x^2(t) + a^2) + p(t)(x(t) + a)\}$$

We calculate the maximum on the right hand side by setting $H_a = -2a + p = 0$. Thus $a = \frac{p}{2}$, and so

$$\alpha(t) = \frac{p(t)}{2}.$$

The dyamical equations are therefore

(ODE) $$\dot{x}(t) = x(t) + \frac{p(t)}{2}$$

and

(ADJ) $$\dot{p}(t) = -H_x = 2x(t) - p(t).$$

Moreover $x(0) = x^0$, and the terminal condition is

(T) $$p(T) = 0.$$

Using the Maximum Principle

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1/2 \\ 2 & -1 \end{pmatrix}}_{M} \begin{pmatrix} x \\ p \end{pmatrix},$$

the general solution of which is

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = e^{tM} \begin{pmatrix} x^0 \\ p^0 \end{pmatrix}.$$

Since we know x^0, the task is to choose p^0 so that $p(T) = 0$.

Feedback controls. An elegant way to do so is to try to find optimal control in linear *feedback* form; that is, to look for a function $c(\cdot) : [0, T] \to \mathbb{R}$ for which

$$\alpha(t) = c(t)\ x(t).$$

We henceforth suppose that an optimal feedback control of this form exists, and attempt to calulate $c(\cdot)$. Now

$$\frac{p(t)}{2} = \alpha(t) = c(t)x(t);$$

whence $c(t) = \frac{p(t)}{2x(t)}$. Define now

$$d(t) := \frac{p(t)}{x(t)};$$

so that $c(t) = \frac{d(t)}{2}$.

Riccati Equation

We will next discover a differential equation that $d(\cdot)$ satisfies. Compute

$$\dot{d} = \frac{\dot{p}}{x} - \frac{p\dot{x}}{x^2},$$

and recall that

$$\begin{cases} \dot{x} = x + \frac{p}{2} \\ \dot{p} = 2x - p. \end{cases}$$

Therefore

$$\dot{d} = \frac{2x - p}{x} - \frac{p}{x^2}\left(x + \frac{p}{2}\right) = 2 - d - d\left(1 + \frac{d}{2}\right) = 2 - 2d - \frac{d^2}{2}.$$

Since $p(T) = 0$, the terminal condition is $d(T) = 0$.

So we have obtained a nonlinear first–order ODE in $p(\cdot)$ with a terminal boundary condition:

$$\text{(R)} \qquad \begin{cases} \dot{d} = 2 - 2d - \frac{1}{2}d^2 & (0 \le t < T) \\ d(T) = 0. \end{cases}$$

This is called the *Riccati equation*.

In summary so far, to solve our linear–quadratic regulator problem, we need to first solve the Riccati equation (R) and then set

$$\alpha(t) = \frac{1}{2}d(t)x(t).$$

Solving the Riccati Equation

- Convert (R) into a second–order, linear ODE

$$d(t) = \frac{2\dot{b}(t)}{b(t)}$$

for a function $b(\cdot)$ to be found. What equation does $b(\cdot)$ solve? We compute

$$\dot{d} = \frac{2\ddot{b}}{b} - \frac{2(\dot{b})^2}{b^2} = \frac{2\ddot{b}}{b} - \frac{d^2}{2}.$$

Hence (R) gives

$$\frac{2\ddot{b}}{b} = \dot{d} + \frac{d^2}{2} = 2 - 2d = 2 - 2\frac{2\dot{b}}{b};$$

and consequently

$$\begin{cases} \ddot{b} = b - 2\dot{b} \qquad (0 \leq t < T) \\ \dot{b}(T) = 0, \ b(T) = 1. \end{cases}$$

This is a terminal-value problem for second–order linear ODE, which we can solve by standard techniques. We then set $d = \frac{2\dot{b}}{b}$, to derive the solution of the Riccati equation (R).

Dynamic Programming

- "it is sometimes easier to solve a problem by embedding it in a larger class of problems and then solving the larger class all at once."

$$(5.1) \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s)) & (t \le s \le T) \\ \mathbf{x}(t) = x. \end{cases}$$

with

$$(5.2) \qquad P_{x,t}[\boldsymbol{\alpha}(\cdot)] = \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s))\, ds + g(\mathbf{x}(T)).$$

Consider the above problems for all choices of starting times $0 \le t \le T$ and all initial points $x \in \mathbb{R}^n$.

DEFINITION. For $x \in \mathbb{R}^n$, $0 \le t \le T$, define the *value function* $v(x,t)$ to be the greatest payoff possible if we start at $x \in \mathbb{R}^n$ at time t. In other words,

$$(5.3) \qquad v(x,t) := \sup_{\boldsymbol{\alpha}(\cdot) \in \mathcal{A}} P_{x,t}[\boldsymbol{\alpha}(\cdot)] \qquad (x \in \mathbb{R}^n, 0 \le t \le T).$$

Notice then that

$$(5.4) \qquad v(x,T) = g(x) \qquad (x \in \mathbb{R}^n).$$

Hamilton-Jacobi-Bellman Equation

- "it's better to be smart from the beginning, than to be stupid for a time and then become smart". choice of life

THEOREM 5.1 (HAMILTON-JACOBI-BELLMAN EQUATION). *Assume that the value function v is a C^1 function of the variables (x,t). Then v solves the nonlinear partial differential equation*

$$(HJB) \quad v_t(x,t) + \max_{a \in A} \{\mathbf{f}(x,a) \cdot \nabla_x v(x,t) + r(x,a)\} = 0 \qquad (x \in \mathbb{R}^n, 0 \le t < T),$$

with the terminal condition

$$v(x,T) = g(x) \qquad (x \in \mathbb{R}^n).$$

Backward induction: change to a sequence of constrained optimization

REMARK. We call (HJB) the *Hamilton–Jacobi–Bellman equation*, and can rewrite it as

$$(HJB) \qquad v_t(x,t) + H(x, \nabla_x v) = 0 \qquad (x \in \mathbb{R}^n, 0 \le t < T),$$

for the *partial differential equations Hamiltonian*

$$H(x,p) := \max_{a \in A} H(x,p,a) = \max_{a \in A} \{\mathbf{f}(x,a) \cdot p + r(x,a)\}$$

where $x, p \in \mathbb{R}^n$. □

Dynamic Programming Method

Step 1: Solve the Hamilton–Jacobi–Bellman equation, and thereby compute the value function v.

Step 2: Use the value function v and the Hamilton–Jacobi–Bellman PDE to design an optimal feedback control $\boldsymbol{\alpha}^*(\cdot)$, as follows. Define for each point $x \in \mathbb{R}^n$ and each time $0 \leq t \leq T$,

$$\boldsymbol{\alpha}(x,t) = a \in A$$

to be a parameter value where the maximum in (HJB) is attained. In other words, we select $\boldsymbol{\alpha}(x,t)$ so that

$$v_t(x,t) + \mathbf{f}(x, \boldsymbol{\alpha}(x,t)) \cdot \nabla_x v(x,t) + r(x, \boldsymbol{\alpha}(x,t)) = 0.$$

Next we solve the following ODE, assuming $\boldsymbol{\alpha}(\cdot, t)$ is sufficiently regular to let us do so:

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}^*(s) = \mathbf{f}(\mathbf{x}^*(s), \boldsymbol{\alpha}(\mathbf{x}^*(s), s)) & (t \leq s \leq T) \\ \mathbf{x}(t) = x. \end{cases}$$

Finally, define the *feedback control*

$$\boldsymbol{\alpha}^*(s) := \boldsymbol{\alpha}(\mathbf{x}^*(s), s). \tag{5.7}$$

In summary, we design the optimal control this way: If the state of system is x at time t, use the control which at time t takes on the parameter value $a \in A$ such that the minimum in (HJB) is obtained.

Example: General Linear Quadratic Regulator

For this important problem, we are given matrices $M, B, D \in \mathbb{M}^{n\times n}$, $N \in \mathbb{M}^{n\times m}$, $C \in \mathbb{M}^{m\times m}$; and assume

$$B, C, D \text{ are symmetric and nonnegative definite,}$$

and

$$C \text{ is invertible.}$$

We take the linear dynamics

$$\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(s) = M\mathbf{x}(s) + N\boldsymbol{\alpha}(s) & (t \le s \le T) \\ \mathbf{x}(t) = x, \end{cases}$$

for which we want to *minimize* the quadratic cost functional

$$\int_t^T x(s)^T B\mathbf{x}(s) + \boldsymbol{\alpha}(s)^T C\boldsymbol{\alpha}(s)\, ds + \mathbf{x}(T)^T D\mathbf{x}(T).$$

So we must *maximize* the payoff

$$\text{(P)} \qquad P_{x,t}[\boldsymbol{\alpha}(\cdot)] = -\int_t^T x(s)^T B\mathbf{x}(s) + \boldsymbol{\alpha}(s)^T C\boldsymbol{\alpha}(s)\, ds - \mathbf{x}(T)^T D\mathbf{x}(T).$$

The control values are unconstrained, meaning that the control parameter values can range over all of $A = \mathbb{R}^m$.

HJB

We will solve by dynamic programming the problem of designing an optimal control. To carry out this plan, we first compute the Hamilton-Jacobi-Bellman equation

$$v_t + \max_{a \in \mathbb{R}^m} \{\mathbf{f} \cdot \nabla_x v + r\} = 0,$$

where

$$\begin{cases} \mathbf{f} = Mx + Na \\ r = -x^T Bx - a^T Ca \\ g = -x^T Dx. \end{cases}$$

Rewrite:

(HJB) $$v_t + \max_{a \in \mathbb{R}^m} \{(\nabla v)^T Na - a^T Ca\} + (\nabla v)^T Mx - x^T Bx = 0.$$

We also have the terminal condition

$$v(x, T) = -x^T Dx$$

Minimization

Minimization. For what value of the control parameter a is the minimum attained? To understand this, we define $Q(a) := (\nabla v)^T N a - a^T C a$, and determine where Q has a minimum by computing the partial derivatives Q_{a_j} for $j = 1, \ldots, m$ and setting them equal to 0. This gives the identitites

$$Q_{a_j} = \sum_{i=1}^{n} v_{x_i} n_{ij} - 2a_i c_{ij} = 0.$$

Therefore $(\nabla v)^T N = 2a^T C$, and then $2C^T a = N^T \nabla v$. But $C^T = C$. Therefore

$$a = \frac{1}{2} C^{-1} N^T \nabla_x v.$$

This is the formula for the optimal feedback control: It will be very useful once we compute the value function v.

Minimization

Finding the value function. We insert our formula $a = \frac{1}{2}C^{-1}N^T\nabla v$ into (HJB), and this PDE then reads

$$\text{(HJB)} \qquad \begin{cases} v_t + \frac{1}{4}(\nabla v)^T N C^{-1} N^T \nabla v + (\nabla v)^T M x - x^T B x = 0 \\ v(x,T) = -x^T D x \end{cases}$$

Our next move is to guess the form of the solution, namely

$$v(x,t) = x^T K(t) x,$$

provided the *symmetric* $n \times n$-matrix valued function $K(\cdot)$ is properly selected. Will this guess work?

Now, since $x^T K(T) x = -v(x,T) = x^T D x$, we must have the terminal condition that

$$K(T) = -D.$$

Minimization

Next, compute that

$$v_t = x^T \dot{K}(t)x, \quad \nabla_x v = 2K(t)x.$$

We insert our guess $v = x^T K(t)x$ into (HJB), and discover that

$$x^T\{\dot{K}(t) + K(t)NC^{-1}N^T K(t) + 2K(t)M - B\}x = 0.$$

Look at the expression

$$\begin{aligned} 2x^T KMx &= x^T KMx + [x^T KMx]^T \\ &= x^T KMx + x^T M^T Kx. \end{aligned}$$

Then

$$x^T\{\dot{K} + KNC^{-1}N^T K + KM + M^T K - B\}x = 0.$$

This identity will hold if $K(\cdot)$ satisfies the *matrix Riccati equation*

$$\text{(R)} \quad \begin{cases} \dot{K}(t) + K(t)NC^{-1}N^T K(t) + K(t)M + M^T K(t) - B = 0 & (0 \le t < T) \\ K(T) = -D \end{cases}$$

In summary, if we can solve the Riccati equation (R), we can construct an optimal feedback control

$$\boldsymbol{\alpha}^*(t) = C^{-1}N^T K(t)\mathbf{x}(t)$$

Relation between DP & Maximum Principle

- Maximum principle starts from 0 to T
- DP starts from t to T
- Costate p at time t is the gradient

The next theorem demonstrates that the costate in the Pontryagin Maximum Principle is in fact the gradient in x of the value function v, taken along an optimal trajectory:

THEOREM 5.3 (COSTATES AND GRADIENTS). *Assume* $\boldsymbol{\alpha}^*(\cdot)$, $\mathbf{x}^*(\cdot)$ *solve the control problem* (ODE), (P).

If the value function v *is* C^2, *then the costate* $\mathbf{p}^*(\cdot)$ *occuring in the Maximum Principle is given by*

$$\mathbf{p}^*(s) = \nabla_x v(\mathbf{x}^*(s), s) \qquad (t \leq s \leq T).$$

Introduction

- Basics
- Controllability
- Linear ODE: Bang-bang control
- Linear time optimal control
- Pontryagin's maximum principle
- Dynamic programming
- Dynamic game

Two-person, Zero-sum Differential Game

- **Basic idea:** Two players control the dynamics of some evolving system, where one tries to maximize, and the other tries to minimize, a payoff functional that depends upon the trajectory

A MODEL PROBLEM. Let a time $0 \leq t < T$ be given, along with sets $A \subseteq \mathbb{R}^m$, $B \subseteq \mathbb{R}^l$ and a function $\mathbf{f} : \mathbb{R}^n \times A \times B \to \mathbb{R}^n$.

DEFINITION. A measurable mapping $\boldsymbol{\alpha}(\cdot) : [t, T] \to A$ is a control for player I (starting at time t). A measurable mapping $\boldsymbol{\beta}(\cdot) : [t, T] \to B$ is a control for player II.

Corresponding to each pair of controls, we have corresponding dynamics:

$$
\text{(ODE)} \qquad \begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \boldsymbol{\alpha}(s), \boldsymbol{\beta}(s)) & (t \leq s \leq T) \\ \mathbf{x}(t) = x, \end{cases}
$$

the initial point $x \in \mathbb{R}^n$ being given.

Two-person, Zero-sum Differential Game

DEFINITION. The *payoff* of the game is

$$(\text{P}) \qquad P_{x,t}[\boldsymbol{\alpha}(\cdot), \boldsymbol{\beta}(\cdot)] := \int_t^T r(\mathbf{x}(s), \boldsymbol{\alpha}(s), \boldsymbol{\beta}(s))\, ds + g(\mathbf{x}(T)).$$

Player I, whose control is $\boldsymbol{\alpha}(\cdot)$, wants to *maximize* the payoff functional $P[\cdot]$. Player II has the control $\boldsymbol{\beta}(\cdot)$ and wants to *minimize* $P[\cdot]$. This is a *two–person, zero–sum differential game.*

DEFINITION. The sets of *controls* for the game of the game are

$$\begin{aligned} \mathcal{A}(t) &:= \{\boldsymbol{\alpha}(\cdot) : [t,T] \to A, \boldsymbol{\alpha}(\cdot) \text{ measurable}\} \\ \mathcal{B}(t) &:= \{\boldsymbol{\beta}(\cdot) : [t,T] \to B, \boldsymbol{\beta}(\cdot) \text{ measurable}\}. \end{aligned}$$

Strategies

- **Idea:** One player will select in advance, not his control, but rather his responses to all possible controls that could be selected by his opponent

DEFINITIONS. (i) A mapping $\Phi : B(t) \to A(t)$ is called a *strategy for player I* if for all times $t \leq s \leq T$,

$$\beta(\tau) \equiv \hat{\beta}(\tau) \quad \text{for } t \leq \tau \leq s$$

implies

(6.1) $$\Phi[\beta](\tau) \equiv \Phi[\hat{\beta}](\tau) \quad \text{for } t \leq \tau \leq s.$$

We can think of $\Phi[\beta]$ as the response of player I to player II's selection of control $\beta(\cdot)$. Condition (6.1) expresses that player I cannot foresee the future.

(ii) A *strategy for II* is mapping $\Psi : M(t) \to N(t)$ such that for all times $t \leq s \leq T$,

$$\alpha(\tau) \equiv \hat{\alpha}(\tau) \quad \text{for } t \leq \tau \leq s$$

implies

$$\Psi[\alpha](\tau) \equiv \Psi[\hat{\alpha}](\tau) \quad \text{for } t \leq \tau \leq s.$$

Value Functions

DEFINITION. The sets of strategies are

$$\mathcal{A}(t) := \text{ strategies for player } I \text{ (starting at } t)$$
$$\mathcal{B}(t) := \text{ strategies for player } II \text{ (starting at } t).$$

DEFINITION. The *lower value function* is

$$v(x,t) := \inf_{\Psi \in \mathcal{B}(t)} \sup_{\boldsymbol{\alpha}(\cdot) \in A(t)} P_{x,t}[\boldsymbol{\alpha}(\cdot), \Psi[\boldsymbol{\alpha}](\cdot)], \tag{6.2}$$

and the *upper value function* is

$$u(x,t) := \sup_{\Phi \in \mathcal{A}} \inf_{\boldsymbol{\beta}(\cdot) \in B(t)} P_{x,t}[\Phi[\boldsymbol{\beta}](\cdot), \boldsymbol{\beta}(\cdot)]. \tag{6.3}$$

Dynamic Programming, Isaacs' Equations

THEOREM 6.1 (PDE FOR THE UPPER AND LOWER VALUE FUNCTIONS). *Assume u, v are continuously differentiable. Then u solves the upper Isaacs' equation*

$$(6.4)\qquad \begin{cases} u_t + \min_{b \in B} \max_{a \in A} \{\mathbf{f}(x,a,b) \cdot \nabla_x u(x,t) + r(x,a,b)\} = 0 \\ u(x,T) = g(x), \end{cases}$$

and v solves the lower Isaacs' equation

$$(6.5)\qquad \begin{cases} v_t + \max_{a \in A} \min_{b \in B} \{\mathbf{f}(x,a,b) \cdot \nabla_x v(x,t) + r(x,a,b)\} = 0 \\ v(x,T) = g(x). \end{cases}$$

Isaacs' equations are analogs of Hamilton–Jacobi–Bellman equation in two–person, zero–sum control theory. We can rewrite these in the forms

$$u_t + H^+(x, \nabla_x u) = 0$$

for the *upper PDE Hamiltonian*

$$H^+(x,p) := \min_{b \in B} \max_{a \in A} \{\mathbf{f}(x,a,b) \cdot p + r(x,a,b)\};$$

and

$$v_t + H^-(x, \nabla_x v) = 0$$

for the *lower PDE Hamiltonian*

$$H^-(x,p) := \max_{a \in A} \min_{b \in B} \{\mathbf{f}(x,a,b) \cdot p + r(x,a,b)\}.$$

Dynamic Programming, Isaacs' Equations

(6.6) $$\max_{a\in A}\min_{b\in B}\{\mathbf{f}(\cdots)\cdot p + r(\cdots)\} = \min_{b\in B}\max_{a\in A}\{\mathbf{f}(\cdots)\cdot p + r(\cdots)\},$$

for all p, x, we say the game satisfies the *minimax condition*, also called *Isaacs' condition*. In this case it turns out that $u \equiv v$ and we say the game has *value*.

(ii) As in dynamic programming from control theory, if (6.6) holds, we can solve Isaacs' for $u \equiv v$ and then, at least in principle, design optimal controls for I and II.

(iii) To say that $\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)$ are optimal means that the pair $(\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot))$ is a *saddle point* for $P_{x,t}$. This means

(6.7) $$P_{x,t}[\boldsymbol{\alpha}(\cdot), \boldsymbol{\beta}^*(\cdot)] \leq P_{x,t}[\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)] \leq P_{x,t}[\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}(\cdot)]$$

for all controls $\boldsymbol{\alpha}(\cdot), \boldsymbol{\beta}(\cdot)$. Player I will select $\boldsymbol{\alpha}^*(\cdot)$ because he is afraid II will play $\boldsymbol{\beta}^*(\cdot)$. Player II will play $\boldsymbol{\beta}^*(\cdot)$, because she is afraid I will play $\boldsymbol{\alpha}^*(\cdot)$. □

Pontryagin's Maximum Principle

Assume the minimax condition (6.6) holds and we design optimal $\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)$ as above. Let $\mathbf{x}^*(\cdot)$ denote the solution of the ODE (6.1), corresponding to our controls $\boldsymbol{\alpha}^*(\cdot), \boldsymbol{\beta}^*(\cdot)$. Then define

$$\mathbf{p}^*(t) := \nabla_x v(\mathbf{x}^*(t), t) = \nabla_x u(\mathbf{x}^*(t), t).$$

It turns out that

$$\text{(ADJ)} \qquad \dot{\mathbf{p}}^*(t) = -\nabla_x H(\mathbf{x}^*(t), \mathbf{p}^*(t), \boldsymbol{\alpha}^*(t), \boldsymbol{\beta}^*(t))$$

for the *game-theory Hamiltonian*

$$H(x, p, a, b) := \mathbf{f}(x, a, b) \cdot p + r(x, a, b).$$

Non-cooperative Differential Game

- Optimization problem for each player can be formulated as the optimal control problem
- The dynamics of state variable and payoff of each player

$$\dot{x}_i(t) = F\left(x_i(t), \mathbf{a}(t)\right) \qquad J_i = \int_0^T \mathscr{U}_i(x_i(t), \mathbf{x}_{-i}(t), a_i(t), \mathbf{a}_{-i}(t))e^{-\rho t}\mathrm{d}t$$

- For player to play the game, the available information is required
- Three cases of available information

  1. Open-loop information

Non-cooperative Differential Game

2. Feedback information

- *At time t, players are assumed to know the values of state variables at time* $t-\epsilon$ *where* ϵ *is positive and arbitrarily small*
- *The feedback information is defined as:*

$$X(t-\epsilon) = \{x_1(t-\epsilon), x_2(t-\epsilon), \ldots, x_N(t-\epsilon)\}$$

3. Closed-loop information

- The Nash equilibrium is defined as a set of action paths of one player to maximize the payoff, given the other players' behavior

Non-cooperative Differential Game

- To obtain the Nash equilibrium, it is required to solve a dynamic optimization problem

$$\max_{a_i(t)} J_i = \int_0^T \mathscr{U}_i(x_i(t), \mathbf{x}_{-i}(t), a_i(t), \mathbf{a}_{-i}(t)) e^{-\rho t} \mathrm{d}t,$$

$$\text{subject to} \quad \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \dot{x}_i(t) = F\left(x_i(t), \mathbf{a}(t)\right),$$

- The Hamiltonian function

$$\mathscr{H}_i(x_i(t), a_i(t)) = e^{-\rho t}\left(\mathscr{U}_i(x_i(t), \mathbf{x}_{-i}(t), a_i(t), \mathbf{a}_{-i}(t)) + \lambda_i(t) F\left(x_i(t), \mathbf{a}(t)\right)\right)$$

  – where $\lambda_i(t) = \mu_i(t) e^{-\rho t}$ is co-state variable. Co-state variable is considered to be the shadow price of the variation of the state variable.

Non-cooperative Differential Game

- The first order conditions for the open-loop solution

$$\begin{aligned}\frac{\partial \mathcal{H}_i(x_i(t), a_i(t))}{\partial a_i(t)} &= 0, \\ -\frac{\partial \mathcal{H}_i(x_i(t), a_i(t))}{\partial x_i(t)} &= \frac{\partial \lambda_i(t)}{\partial t},\end{aligned}$$

- For the closed-loop solution, the conditions are slightly different

$$\begin{aligned}\frac{\partial \mathcal{H}_i(x_i(t), a_i(t))}{\partial a_i(t)} &= 0, \\ -\frac{\partial \mathcal{H}_i(x_i(t), a_i(t))}{\partial x_i(t)} - \sum_{j\neq i} \frac{\partial \mathcal{H}_i(x_i(t), a_i(t))}{\partial a_j(t)} \frac{\partial a_j(t)}{\partial x_i(t)} &= \frac{\partial \lambda_i(t)}{\partial t}.\end{aligned}$$

- Further reading: Basar’s book

Summary of Dynamic Control

- Dynamic problem formulation
  - ODE and payoff function
- Conditions for controllability
  - Rank of G and eigenvalue of M
- Bang-bang control
- Maximum principle
  - ODE, ADJ, M and P
- Dynamic programming
  - Divide a complicated problem into sequence of sub-problems
  - HJB equations
- Dynamic game: Multiuser case
- Future reading: Stochastic game

Applications in Wireless Networks

Packet Routing

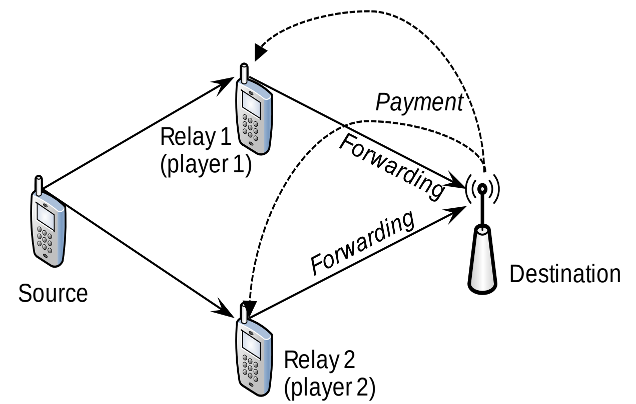
- For routing in the mobile ad hoc network (MANET), the forwarding nodes as the players have incentive from the destination in terms of price to allocate transmission rate to forward packets from source
- A differential game for duopoly competition is applied to model this competitive situation

L. Lin, X. Zhou, L. Du, and X. Miao. Differential game model with coupling constraint for routing in ad hoc networks. In *Proc. of the 5th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2009), pages 3042-3045, September 2009.*

Applications in Wireless Networks

Packet Routing

- There are two forwarding nodes that are considered to be the players in this game



- Destination pays some price to forwarding nodes according to the amount of forwarded data
- Forwarding nodes compete with each other by adjusting the forwarding rate (i.e., action denoted by $a_i(t)$ for player i at time t) to maximize their utilities over time duration of $[0,\infty]$

Applications in Wireless Networks

Packet Routing

- Payment from the destination at time t is denoted by $P(t)$
- Payoff function of player i can be expressed as follows:

$$\boxed{J_i = \int_0^{+\infty} e^{-\rho t} \left(P(t)a_i(t) \underbrace{- c a_i(t) - \frac{1}{2} a_i(t)^2}_{\text{Quadratic cost function}} - g(\mathbf{a}) \right) \mathrm{d}t,}$$

- $P(t)a_i(t)$ is revenue

- $g(\mathbf{a})$ is a cost function given vector $\mathbf{a}$ of actions of players

- For the payment, the following evolution of price (i.e., a differential equation of Tsutsui and Mino) is considered

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = \dot{P}(t) = K(E - a_1(t) - a_2(t) - P(t)),$$

Applications in Wireless Networks

Packet Routing

- Using optimal control approach, feedback Nash equilibrium strategies of this game can be expressed as follows

$$\boxed{a_i^* = P^*(t) - c - K(AP^*(t) - B),}$$

- Iterative approach based on greedy adjustment is proposed to obtain the solution
- Algorithm gradually increases the forwarding rate of the player as long as the payoff is non-decreasing
- If the payoff of one player decreases, the algorithm will allow the other players to adjust the forwarding rate until none of players can gain a higher payoff

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

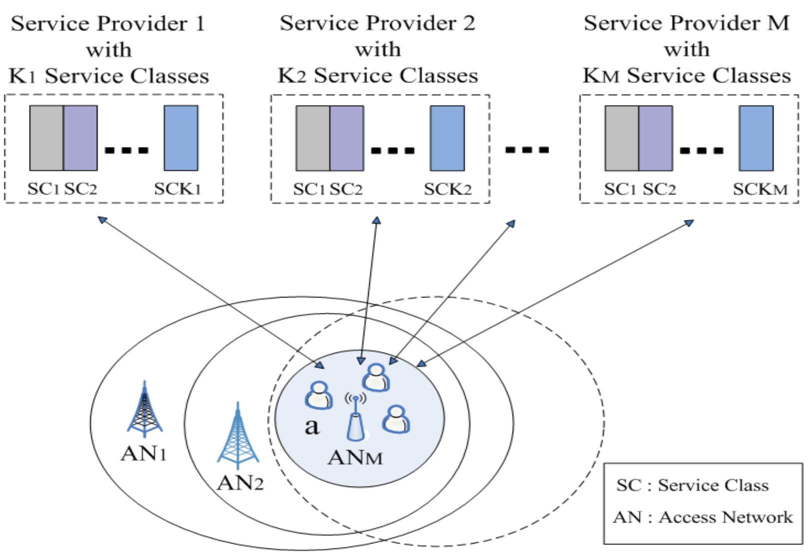
- In heterogeneous wireless network, user can access multiple wireless networks (e.g., 3G, WiFi, WiMAX)
- However, none of the existing works consider the dynamic bandwidth allocation in heterogeneous wireless networks in which the users can change service selection dynamically
- The network systems are naturally dynamic, a steady state of the network may never be reached
- Therefore, the dynamic optimal control is the suitable approach for analyzing the dynamic decision making process

Z. Kun, **D. Niyato**, and P. Wang, "Optimal bandwidth allocation with dynamic service selection in heterogeneous wireless networks," in *Proceedings of IEEE GLOBECOM'10*, Miami FL USA, 6-10 December 2010.

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

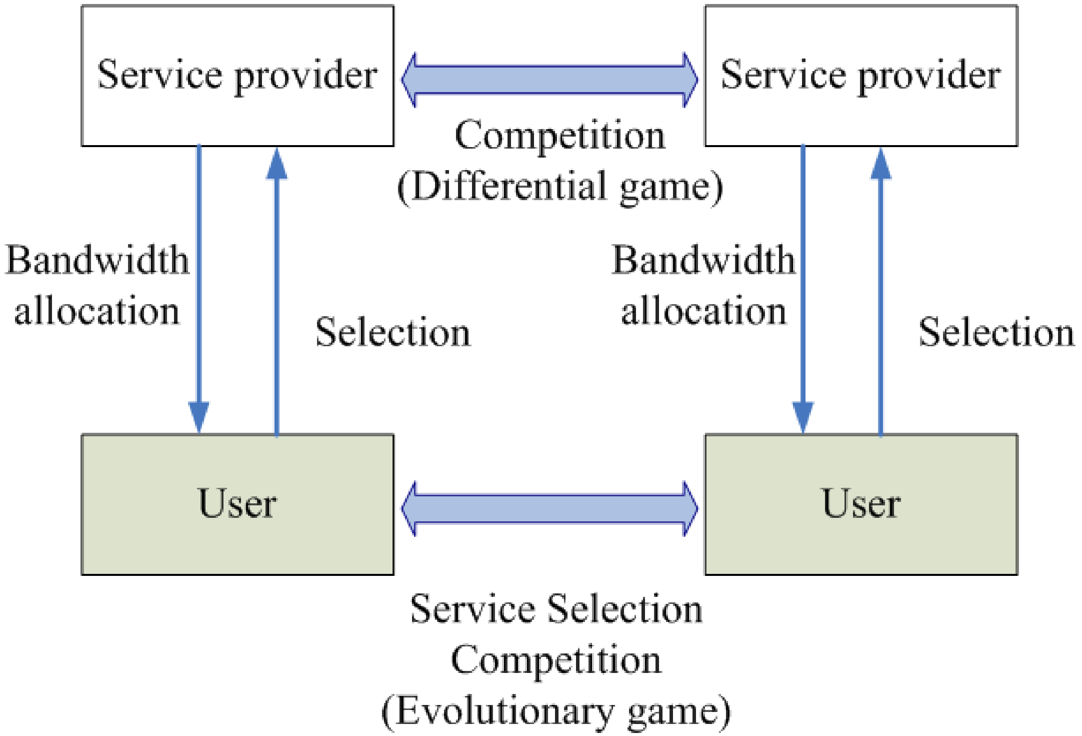
- Designing a dynamic game framework for optimal bandwidth allocation under dynamic service selection
  - For service providers: the profit can be maximized
  - For users: the performance can be maximized under competition



Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Two-level game framework for optimal bandwidth allocation with dynamic service selection



Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Game formulation: Evolution of Service Selection
  - Players: N active users in area a
  - Strategy: The choices of particular service class from certain service providers
  - Payoff: The payoff of user k selecting service class j from service provider i :

$$u(\tau_k^{ij}(t)) = \alpha \tau_k^{ij}(t) = \alpha \frac{B_{ij}(t)}{N(t)x_{ij}(t)},$$

  - The replicator dynamics modeling the service selection:

$$\frac{\partial x_{ij}(t)}{\partial t} = \dot{x}_{ij}(t) = \delta x_{ij}(t)\left(u(\tau_k^{ij}(t)) - \bar{u}(t)\right), \qquad \sum_{i=1}^{M}\sum_{j=1}^{K_i} x_{ij}(t) = 1.$$

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Game formulation: Dynamic Bandwidth Allocation
  - Players: *M* service providers in area a
  - Control strategies: The control strategy of player *i* denoted by

$$\vec{\gamma}_i(t) = [\ \gamma_{i1}(t)\ \cdots\ \gamma_{ij}(t)\ \cdots\ \gamma_{iK_i}(t)\ ]^T \in R_+^{K_i}.$$

  - Open-loop vs Closed-loop
  - System state:

$$\vec{\mathbf{x}}(t) = [\ x_{11}(t)\ \cdots\ x_{ij}(t)\ \cdots\ x_{MK_M}(t)\ ]^T.$$

  - The instantaneous payoff:

$$J^i_{\text{ins}}(\vec{\gamma}_i(t), \vec{\gamma}_{-i}(t)) = \sum_{j=1}^{K_i} (P_{ij} N(t) x_{ij}(t) - \theta_j (\gamma_{ij}(t) B_i(t))^2).$$

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Optimal Control Formulation

maximize:

$$J^i(\vec{\gamma}_i(t), \vec{\gamma}_{-i}(t)) = \int_0^{\infty} e^{-\rho t} \sum_{j=1}^{K_i} (P_{ij} N(t) x_{ij}(t) - \theta_j(\gamma_{ij}(t) B_i(t))^2) \mathrm{d}t,$$

subject to:

$$\begin{aligned} \dot{x}_{ij}(t) &= \delta x_{ij}(t) \left(u \left(\frac{B_i(t)\gamma_{ij}(t)}{N(t) x_{ij}(t)} \right) - \overline{u}(t) \right), \\ \vec{\mathrm{x}}(0) &= \vec{\mathrm{x}}_0, \end{aligned}$$

for $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, K_i\}$, where

$$\begin{aligned} \sum_{i=1}^{M} \sum_{j=1}^{K_i} x_{ij}(t) &= 1, \quad x_{ij}(t) \in [0, 1], \\ \sum_{j=1}^{K_i} \gamma_{ij}(t) &= 1, \quad \gamma_{ij}(t) \in [0, 1], t \in [0, +\infty), \end{aligned}$$

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Open-loop Nash equilibrium

Definition 1: A bandwidth allocation control path $\vec{\gamma}_i^*(t)$ is optimal for service provider i if the inequality condition $J^i(\vec{\gamma}_i^*(t), \vec{\gamma}_{-i}(t)) \geq J^i(\vec{\gamma}_i(t), \vec{\gamma}_{-i}(t))$ holds for all feasible control paths $\vec{\gamma}_i(t)$ in the noncooperative bandwidth allocation differential game.

Definition 2: Denote $\vec{\gamma}_i(t)$ the open-loop bandwidth allocation strategy of service provider i. The strategy profile $\Phi = \{\vec{\gamma}_i^*(t), \vec{\gamma}_{-i}^*(t)\}$ is an open-loop Nash equilibrium if for each service provider $i \in \{1, 2, \ldots, M\}$, $\vec{\gamma}_i^*(t)$ is an optimal control path given other service providers' control strategies $\vec{\gamma}_{-i}^*(t)$.

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Pontryagin's Maximum Principle for Nash Equilibrium

  A strategy profile $(\overrightarrow{\gamma *_i}(t), \vec{\gamma}*_{-i}(t))$ is Nash Equilibrium if there exists $\lambda *_{ij}(t)$ for every optimal control path such that the following conditions are satisfied

  – The maximum condition holds for all players

$$H_i(\vec{\mathbf{x}}(t), \vec{\gamma}*_i(t), \overrightarrow{\gamma *}_{-i}(t), \lambda *_{ij}(t), t) = H_i^*(\vec{\mathbf{x}}(t), \lambda *_{ij}(t), t)$$

  – Adjoint equation $\dot{\lambda} *_{ij}(t) = \rho \lambda *_{ij}(t) - \dfrac{\partial H_i^*(\vec{\mathbf{x}}(t), \lambda *_{ij}(t), t)}{\partial x_{ij}(t)}$ holds for all i, j

  – The constraints and boundary conditions are satisfied

  – $H_i^*(\vec{\mathbf{x}}(t), \lambda *_{ij}(t), t)$ is concave and continuously differentiable

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Cooperative Bandwidth Allocation
  - Maximize:

$$J(\vec{\gamma}_i(t), \vec{\gamma}_{-i}(t)) = \int_0^\infty e^{-\rho t} \sum_{i=1}^{M} \sum_{j=1}^{K_i} (P_{ij} N(t) x_{ij}(t) - \theta_j (\gamma_{ij}(t) B_i(t))^2) \mathrm{d}t.$$

  - The Hamiltonian function:

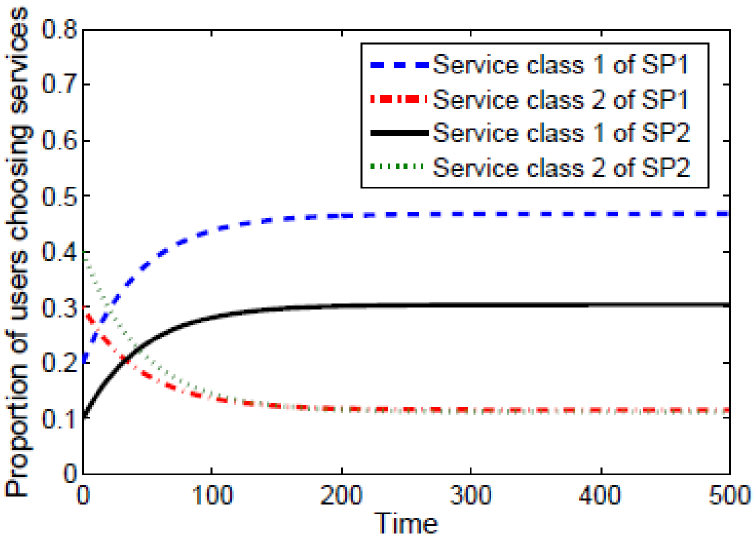
$$H_i^c(\vec{\mathbf{x}}(t), \vec{\gamma}_i(t), \vec{\gamma}_{-i}(t), \lambda_{ij}^c(t), t) = \sum_{i=1}^{M} \sum_{j=1}^{K_i} (P_{ij} N(t) x_{ij}(t) - \theta_j (\gamma_{ij}(t) B_i(t))^2) + \sum_{i=1}^{M} \sum_{j=1}^{K_i} \lambda_{ij}^c(t) \delta x_{ij}(t) \left(u\left(\frac{B_i(t)\gamma_{ij}(t)}{N(t) x_{ij}(t)} \right) - \bar{u} \right),$$

  - Observation: In the non-cooperative bandwidth allocation differential game, the selfish behavior of service providers can also maximize the social welfare

Applications in Wireless Networks

Dynamic Bandwidth Allocation with Dynamic Service Selection in Heterogeneous Wireless Networks

- Convergence
  - The strategy adaption trajectory of the lower level service selection evolutionary game from the initial selection distribution



Dynamics of service selection.

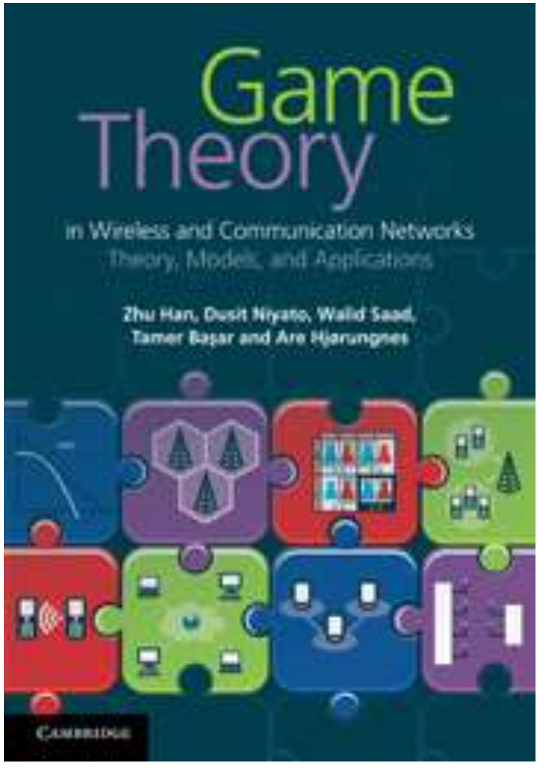
Summary

- The basics of differential game have been discussed
- Two applications of differential game in wireless network, i.e., routing and bandwidth allocation have been presented
- Differential game can be used in other applications (e.g., cognitive radio) which are open to exploration

Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 6
Auction Theory

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of Lecture Notes

- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, Chapter 2, book 2
- Contract Theory, Lecture 8, Chapter 3, book 2
- Learning in Game, Lecture 9, Chapter 6, book 2
- Stochastic Game, Lecture 10, Chapter 4, book 2
- Game with Bounded Rationality, Lecture 11, Chapter 5, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 12, Chapter 7, book 2
- Zero Determinant Strategy, Lecture 13, Chapter 8, book 2
- Mean Field Game, Lecture 14, book 2
- Network Economy, Lecture 15, book 2

Outline

- Introduction and auction basics
- Mechanism design
- Special auctions
  - VCG auction
  - Share auction
  - Double auction
  - Sequential first-price auction
  - AGV auction
  - Ascending clock auction
  - (Reverse) combinatorial auction
- Auction theory in spectrum sharing
- Summary

Introduction

- A market mechanism in which an object, service, or set of objects, is exchanged on the basis of bids submitted by participants.
- An auction provides a specific set of rules that will govern the sale or purchase (procurement auction) of an object to the submitter of the most favorable bid.
- Typical auctions
  - First-price auction
  - Second-price auction
  - English auction
  - Dutch auction
  - Japanese auction
  - All-pay actions
  - Unique bid auction
  - Generalized second-price auction

Properties of Auctions

- Allocative efficiency means that in all these auctions the highest bidder always wins (i.e., there are no reserve prices).
- It is desirable for an auction to be computationally efficient.
- Revenue Equivalence Theorem: Any two auctions such that:
  - *The bidder with the highest value wins*
  - *The bidder with the lowest value expects zero profit*
  - *Bidders are risk-neutral 1*
  - *Value distributions are strictly increasing and atomless*

  have the same revenue and also the same expected profit for each bidder. The theorem can help find some equilibrium strategy.

Mechanism Design

- Definition of Mechanism
- Outcome set: Ω.
- Players $i \in \mathcal{I}$, where $\mathcal{I}$ is the set of the players of size $|\mathcal{I}| = N$, with preference types $\theta_i \in \Theta_i$.
- Utility $u_i(o, \theta_i)$, over outcome $o \in \Omega$.
- Mechanism $M = (S, g)$ defines:
  - a strategy space $S^N = S_1 \times \cdots \times S_N$, s.t. player i chooses a strategy $s_i(\theta_i) \in S_i$ with $s_i : \Theta_i \to S_i$;
  - an outcome function $g : S^N \to \Omega$, s.t. outcome $g(s_i(\theta_1), \ldots, s_N(\theta_N))$ is implemented given strategy profile $s = (s_1(\cdot), \ldots, s_N(\cdot))$.
- Game: The utility to player i from strategy profile s, is $u_i(g(s(\theta)), \theta_i)$, which denotes as $u_i(s, \theta_i)$.

Design goal and properties

- The objective of a mechanism $M = (S, g)$ *is to achieve the desired game* outcome

$$g(s_1^*(\theta_1), \ldots, s_N^*(\theta_N)) = f(\theta), \ \forall \theta \in \Theta^N$$

- Desired properties
  - Efficiency: select the outcome that maximizes total utility.
  - Fairness: select the outcome that achieves a certain fairness criterion in utility.
  - Revenue maximization: select the outcome that maximizes revenue to a seller (or more generally, utility to one of the players).
  - Budget-balanced: implement outcomes that have balanced transfers across players.
  - Pareto optimality

Equilibrium concepts

Definition 62 *Nash implementation: Mechanism* $M = (S, g)$ *implements* $f(\theta)$ *in Nash equilibrium if, for all* $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, *where* $s^*(\theta)$ *is a Nash equilibrium, i.e.,*

$$u_i(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i) \geq u_i(s_i'(\theta_i), s_{-i}'(\theta_{-i}), \theta_i), \ \forall i, \forall \theta_i, \forall s_i' \neq s_i^*. \tag{8.7}$$

Definition 63 *Bayes-Nash implementation: With common prior* $F(\theta)$, *mechanism* $M = (S, g)$ *implements* $f(\theta)$ *in Bayes-Nash equilibrium if, for all* $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, *where* $s^*(\theta)$ *is a Bayes-Nash equilibrium; i.e.,*

$$E_{\theta_{-i}}[u_i(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i)] \geq E_{\theta_{-i}}[u_i(s_i'(\theta_i), s_{-i}'(\theta_{-i}), \theta_i)], \ \forall i, \forall \theta_i, \forall s_i' \neq s_i^*. \tag{8.8}$$

Definition 64 *Dominant implementation: Mechanism* $M = (S, g)$ *implements* $f(\theta)$ *in a dominant strategy equilibrium if, for all* $\theta \in \Theta$, $g(s^*(\theta)) = f(\theta)$, *where* $s^*(\theta)$ *is a dominant strategy equilibrium; i.e.,*

$$u_i(s_i^*(\theta_i), s_{-i}^*(\hat{\theta}_{-i}), \theta_i) \geq u_i(s_i'(\theta_i), s_{-i}'(\hat{\theta}_{-i}), \theta_i), \ \forall i, \forall \theta_i, \forall \hat{\theta}_{-i}, \forall s_i' \neq s_i^*. \tag{8.9}$$

Rationality Concepts

Let $\bar{u}_i(\theta_i)$ denote the (expected) utility to player i with type θ_i as its outside option, and recall that $u_i(f(\theta);\theta_i)$ is the equilibrium utility of player i from the mechanism. We have the following three definitions of rationality.

- Ex ante individual rationality: players choose to participate before they know their own types:

$$E_{\theta\in\Theta}[u_i(f(\theta);\theta_i)] \geq E_{\theta_i\in\Theta_i}[\bar{u}_i(\theta_i)]. \tag{8.10}$$

- Interim individual rationality: players can withdraw once they know their own type:

$$E_{\theta_{-i}\in\Theta_{-i}}[u_i(f(\theta,\theta_{-i});\theta_i)] \geq \bar{u}_i(\theta_i). \tag{8.11}$$

- Ex post individual rationality: players can withdraw from the mechanism at the end:

$$u_i(f(\theta);\theta_i) \geq \bar{u}_i(\theta_i). \tag{8.12}$$

DRM, Incentive Compatible and Strategy Proof

A special kind of mechanisms is called the direct-revelation mechanism (DRM), which has a strategy space $S = \Theta$ and which has a player simply reports a type to the mechanism with outcome rule $g : \Theta \to \Omega$. For DRM, we have the following definitions for incentive compatible and strategy proof.

Definition 65 *Incentive compatible: A DRM is (Bayes-)Nash incentive compatible if truth revelation is a (Bayes-)Nash equilibrium; i.e., $s_i^*(\theta_i) = \theta_i$, for all $\theta \in \Theta$.*

Definition 66 *Strategy proof: A DRM is strategy proof if truth revelation is a dominant strategy equilibrium, for all $\theta \in \Theta$.*

Revelation Principle

- For any Bayesian Nash equilibrium there corresponds a Bayesian game with the same equilibrium outcome but in which players truthfully report type.
- Allows one to solve for a Bayesian equilibrium by assuming all players truthfully report type (subject to an incentive-compatibility constraint), which eliminates the need to consider either strategic behavior or lying.
- no matter what the mechanism, a designer can confine attention to equilibria in which players truthfully report type.
- Theorem: *For any mechanism, M, there is a direct, incentive-compatible mechanism with the same outcome.*
  - *Incentive compatible is free*
  - *Fancy mechanisms are unnecessary*

Budget Balance

- Transfer or side payment

Define the outcome space $\mathcal{O} = \mathcal{K} \times \mathbb{R}^N$, such that an outcome rule, $o = (k, t_1, \ldots, t_N)$, defines a choice, $k(s) \in \mathcal{K}$, and a transfer, $t_i(s) \in \mathbb{R}$ from player i to the mechanism, given strategy profile $s \in S$. For example, the utility can be written as

$$u_i(o, \theta_i) = v_i(k, \theta_i) - t_i, \tag{8.15}$$

where $v_i(k, \theta_i)$ is the value of player i and t_i is the payment (transfer) to the auctioneer.

- Budget balance

Definition 67 *Budget balance introduces constraints over the total transfers made from the players to the mechanism. Let $s^*(\theta)$ denote the equilibrium strategy of a mechanism. We have*

1. *Weak budget balance*
   (a) *ex post:* $\sum_t t_i(s^*(\theta)) \geq 0, \forall\theta;$
   (b) *ex ante:* $E_{\theta\in\Theta}[\sum_t t_i(s^*(\theta))] \geq 0.$

2. *Strong budget balance*
   (a) *ex post:* $\sum_t t_i(s^*(\theta)) = 0, \forall\theta;$
   (b) *ex ante:* $E_{\theta\in\Theta}[\sum_t t_i(s^*(\theta))] = 0.$

Efficiency

Definition 68 *A choice rule, $k^* : \Theta \rightarrow \mathcal{K}$, is (ex post) efficient if for all $\theta \in \Theta$, $k^*(\theta)$ maximizes the sum of individual value functions $\sum_{k \in \mathcal{K}} v_i(k, \theta_i)$.*

Unfortunately, according to Green-Laffont impossibility theorem [170], if Θ allows all valuation functions from $\mathcal{K}$ to $\mathbb{R}$, then no mechanism can implement an efficient and ex post budget-balanced in dominant strategy. So we can either (1) restrict space of preferences, (2) drop budget-balance, (3) drop efficiency, or (4) drop dominant strategy.

Groves Mechanisms

Definition 69 *A Groves mechanism,* $M = (\Theta, k, t_1, \ldots, t_N)$*, is defined with choice rule,*

$$k^*(\hat{\theta}) = \arg\max_{k \in \mathcal{K}} \sum_i v_i(k, \hat{\theta}_i), \tag{8.16}$$

and transfer rules

$$t_i(\hat{\theta}) = h_i(\hat{\theta}_{-i}) - \sum_{j \neq i} v_j(k^*(\hat{\theta}), \hat{\theta}_j), \tag{8.17}$$

where $h_i(\cdot)$ *is an (arbitrary) function that does not depend on the reported type,* $\hat{\theta}_i$*, of player* i*.*

It has been proved that Groves mechanisms are strategy proof and efficient [171]. Groves mechanisms are unique, in the sense that any mechanism that implements efficient choice, $k^*(\theta)$, in truthful dominant strategy must implement Groves transfers.

Impossibility and Possibility

Theorem 16 Gibbard-Satterthwaite Impossibility Theorem [162, 425]: *If agents have general preferences, there are at least two agents, and at least three different optimal outcomes over the set of all agent preferences, then a social-choice function is dominant-strategy implementable if and only if it is dictatorial (i.e., one (or more) agents always receive one of its most preferred alternatives).*

Theorem 17 Hurwicz Impossibility Theorem [218]: *It is impossible to implement an efficient, budget-balanced, and strategy-proof mechanism in a simple exchange economy* [3] *with quasi-linear preferences.*

Theorem 18 Myerson-Satterthwaite Theorem [345]: *It is impossible to achieve allocative efficiency, budget balance and (interim) individual rationality in a Bayesian-Nash incentive-compatible mechanism, even with quasi-linear utility functions.*

Outline

- Introduction and auction basics
- Mechanism design
- Special auctions
  - VCG auction
  - Share auction
  - Double auction
  - Sequential first-price auction
  - AGV auction
  - Ascending clock auction
  - Reverse combinatorial auction
- Summary

VCG Auction

- Vickrey auction is a type of sealed-bid auction, in which bidders (players) submit written bids without knowing the bid of the other people in the auction. The highest bidder wins, but the price paid is the second-highest bid.

Definition 70 *The VCG mechanism implements efficient outcome,* $k^* = \max_k \sum_j v_j(k, \hat{\theta}_j)$, *and computes transfers*

$$t_i(\hat{\theta}) = \sum_{j \neq i} v_j(k^{-i}, \hat{\theta}_j) - \sum_{j \neq i} v_j(k^*, \hat{\theta}_j), \qquad (8.18)$$

where $k^{-i} = \max_k \sum_{j \neq i} v_j(k, \hat{\theta}_j)$.

- In other words, the payment equals to the performance loss of all other users because of including user *i* .
- Truthful relevance, ex post efficient, and strategy proof

VCG Example

For example, suppose two apples are being auctioned among three bidders. Bidder A wants one apple and bids \$5 for that apple. Bidder B wants one apple and is willing to pay \$2 for it. Bidder C wants two apples and is willing to pay \$6 to have both of them, but is uninterested in buying only one without the other. First, the outcome of the auction is determined by maximizing bids: the apples go to bidder A and bidder B. Next, to decide payments, the opportunity cost each bidder imposed on the rest of the bidders is considered. Currently, B has a utility of \$2. If bidder A had not been present, C would have won, and had a utility of \$6, so A pays \$6-\$2 = \$4. For the payment of bidder B: currently A has a utility of \$5 and C has a utility of 0. If bidder B had been absent, C would have won and had a utility of \$6, so B pays \$6-\$5 = \$1. The outcome is identical whether or not bidder C participates, so C does not need to pay anything.

Shortcoming of VCG

- It does not allow for price discovery - that is, discovery of the market price if the buyers are unsure of their own valuations - without sequential auctions.
- Sellers may use shill bids to increase profit.
- In iterated Vickrey auctions, the strategy of revealing true valuations is no longer dominant.
- It is vulnerable to collusion by losing bidders.
- It is vulnerable to shill bidding with respect to the buyers.
- It does not necessarily maximize seller revenues; seller revenues may even be zero in VCG auctions. If the purpose of holding the auction is to maximize profit for the seller rather than just allocate resources among buyers, then VCG may be a poor choice.
- The seller's revenues are non-monotonic with regard to the sets of bidders and offers.

Share Auction

- A share auction is concerned with allocating a perfect divisible good among a set of bidders.
- Share Auction Mechanism:

1. The manager announces a reserve bid $\beta \geq 0$, and a price $\pi > 0$.
2. After observing β, π, user i submits a bid $b_i \geq 0$.
3. The resources are allocated to each user i, and its share p_i is proportional to its bid, i.e.,

$$p_i = \frac{\beta}{\sum_i b_i + \beta} P. \tag{8.19}$$

   The resulting performance of user i is γ_i.

   For example, if P is the overall transmitted power, for the interference case, we can have the resulting SINR for user i as

$$\gamma_i(\mathbf{p}) = \frac{p_i h_{ii}}{n_0 + \sum_{j \neq i} p_j h_{ji}}, \tag{8.20}$$

   where h_{ij} is the channel gain for user i to receiver j and n_0 is the noise level.

4. In a share auction, user i pays $C_i = \pi \gamma_i$.

Share Auction

A *bidding profile* is the vector containing the users' bids $\mathbf{b} = (b_1, ..., b_N)$. The *bidding profile of user i's opponents* is defined as $b_{-i} = (b_1, ..., b_{i-1}, b_{i+1}, ..., b_N)$, so that $\mathbf{b} = (b_i; b_{-i})$. Typically, each user i submits a bid b_i to maximize its *surplus function*

$$S_i(b_i; b_{-i}) = U_i(\gamma_i(b_i; b_{-i})) - C_i.$$

These auction mechanisms differ from some previously proposed auction-based network resource allocation schemes (e.g., [234, 312]) in that the bids here are not the same as the payments. Instead, the bids are signals of willingness to pay. The manager can, therefore, influence the NE by choosing β and π. This alleviates the typical inefficiency of the NE, and in some cases allows us to achieve socially optimal solutions.

Under properly chosen price π, the share auction can achieve fair (efficient, respectively) allocation. In a fair allocation, users achieve the performance by some predefined fairness criteria. In an efficient allocation, the total utility of the network is maximized.

Double Auction

- A **double** auction is a process of buying and selling goods when potential buyers submit their bids and potential sellers simultaneously submit their ask prices to an auctioneer, and then an auctioneer chooses some price *p* that clears the market: all the sellers who asked less than p sell and all buyers who bid more than *p* buy at this price *p*.

Consider a double auction with a single buyer and a single seller. Suppose that the valuation of a buyer is *v* and the valuation of a seller is *c* (e.g. the cost of producing the product). And *v*, *c* $\in [0,1]$.Submitted bid of a seller is b_1, and bid of a buyer is b_2. $b_1, b_2 \in [0,1]$. Let $v > c$.

Suppose an auctioneer sets the price:

$p = \dfrac{(b_1 + b_2)}{2}$ if $b_1 \leq b_2$. And if $b_1 > b_2$ trade does not occur.

Consumer surplus of buyer is $u_1 = v - p$ if $b_1 \leq b_2$ and 0 if $b_1 > b_2$

Producer surplus of a seller is $u_2 = p - c$ if $b_1 \leq b_2$ and 0 if $b_1 > b_2$

In a *complete information* (symmetric information) case when the valuations are common knowledge it can be shown that the continuum of pure strategy efficient Nash equilibriums exists with $b_1 = b_2 = p \in [c, v]$.

In an *incomplete information* (asymmetric information) case a buyer and a seller know only their own valuations. Suppose that these valuations are uniformly distributed over the same interval. Then it can be shown that such a game has a Bayesian Nash equilibrium with linear strategies. That is there is an equilibrium when both players' bids are some linear functions of their valuations. It is also the equilibrium that brings the highest expected gains for the players than any other Bayesian Nash equilibrium[1]

Double Auction

- Buyer utility

$$U_i^{(b)} = \sum_{n=1}^{\hat{n}} (p_i^{(b)} - \hat{p}_{i,n})\hat{x}_{i,n},$$

- Seller utility

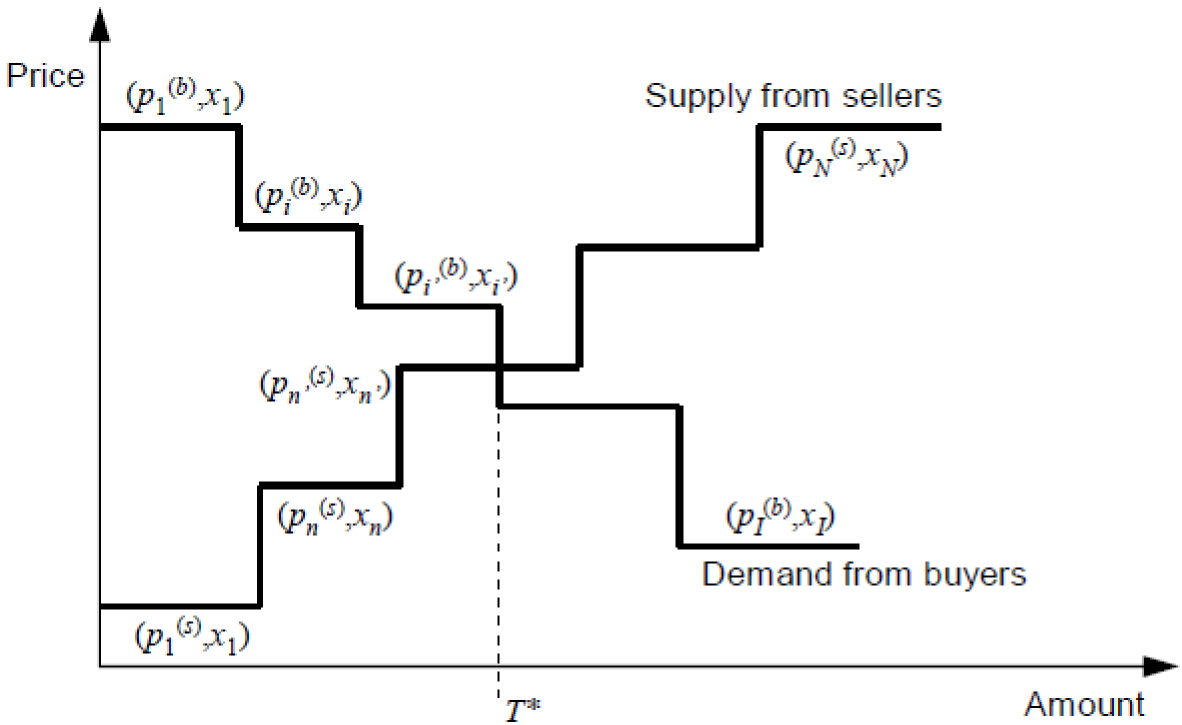
$$U_n^{(s)} = \sum_{i=1}^{\hat{i}} (\hat{p}_{i,n} - p_n^{(s)})\hat{x}_{i,n}.$$

- Optimization

$$\max \quad \sum_{i=1}^{\hat{i}} \sum_{n=1}^{\hat{n}} \hat{x}_{i,n} \left(p_i^{(b)} - p_n^{(s)} \right),$$

$$\text{s.t.} \quad \sum_{i=1}^{\hat{i}} \hat{x}_{i,n} \le y_n, \forall n, \sum_{n=1}^{\hat{n}} \hat{x}_{i,n} \le x_i, \forall i, \hat{x}_{i,n} \ge 0, \forall i, n.$$

Double Auction



Example of Ordered Demand and Supply in a Double Auction.

Sequential First/Second-Price Auction

- In multi-object auctions, multiple objects are put up for auction and multiple winners emerge.
  - substitutable auctions: different objects have same value for specific bidder.
  - non-substitutable auctions: the values of different objects are different
- The winners of the items sequentially decided in multiple rounds. In each round, each bidder can submit only one bid for a specific object and only the winners of the bided items are decided.
- Sequential auctions require less information exchange among the auctioneer and bidders, and provide high flexibility for the bidders and high revenue for the auctioneer.

Sequential First/Second-Price Auction

- The sequential first price auction is a situation in which the *M object are sold to N > M bidders using a series of first-price sealed bid* auctions.
  - Specifically, in each round, each bidder bids for a specific item, and for each item that is bided by at least one bidder, the item is obtained by the bidder with the highest bid at the price of its submitted bid.
- The only difference for the sequential second price auction is that, in each round, the winner obtains the item at the price of the second highest bid, instead of its submitted bid.

AGV Auction

- AGV(Arrow-d' Aspremont-Gerard-Varet) mechanism is an extension of the Groves mechanism,
  - Incentive Compatibility, Individual Rationality and Budget Balance
  - "Expected form" of the Groves mechanisms
  - The allocation rule is the same as VCG.

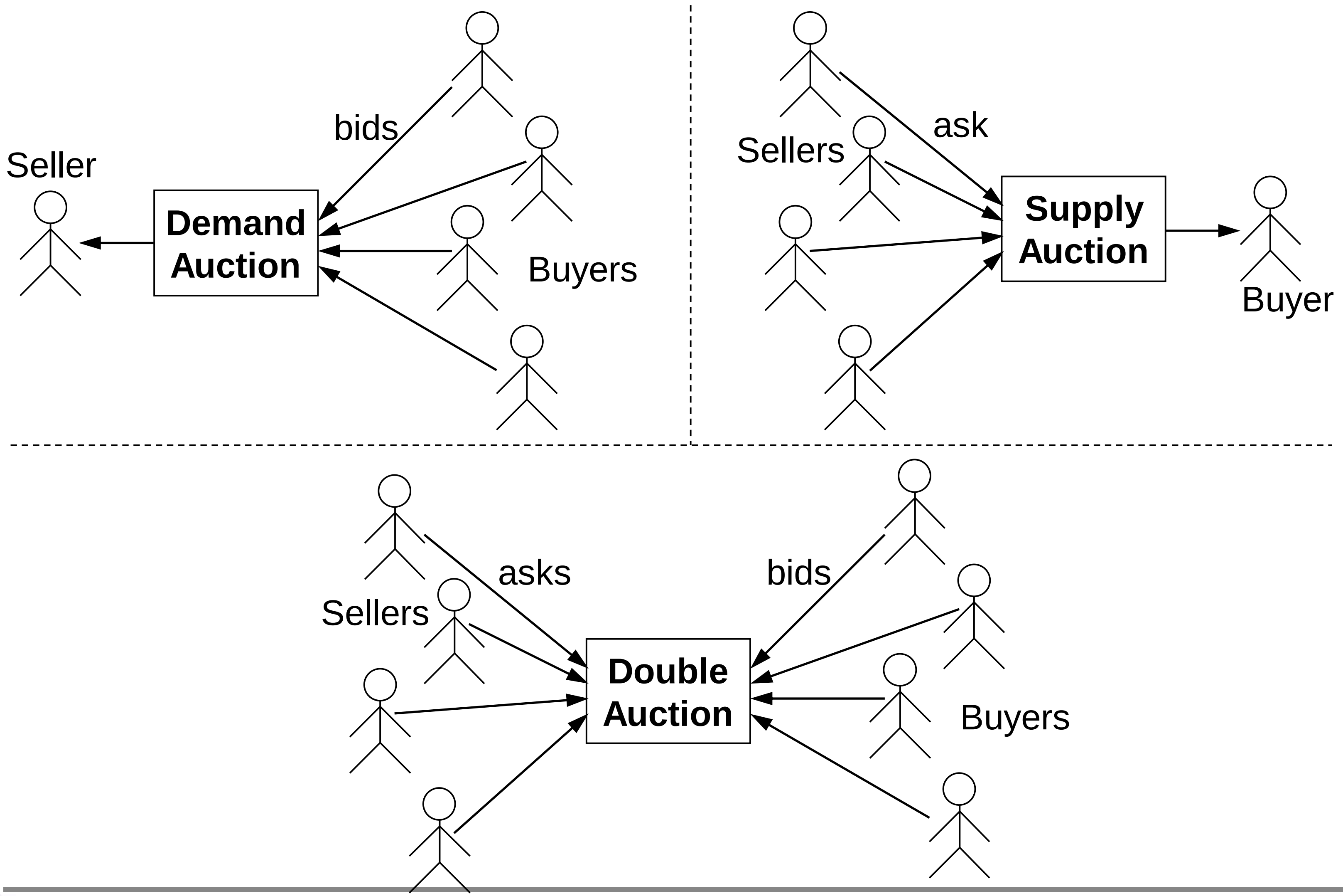
Ascending Clock Auction

- During the auction, the auctioneer first announces a initial price, then the bidders report to the auctioneer their demands at that price, and the auctioneer raises the price until the total demands meet the supply.
  - Distributed approach
  - Single object pay-as-bid ascending clock auction, in which the good is sold as a single object and the bidder can only bid 0 or *p*
  - Traditional ascending clock auction, where each bidder is allowed to bid between 0 and p at every iteration. Distributed dual-based optimization approach for Network Utility Maximization problem

(Reverse) Combinatorial Auction

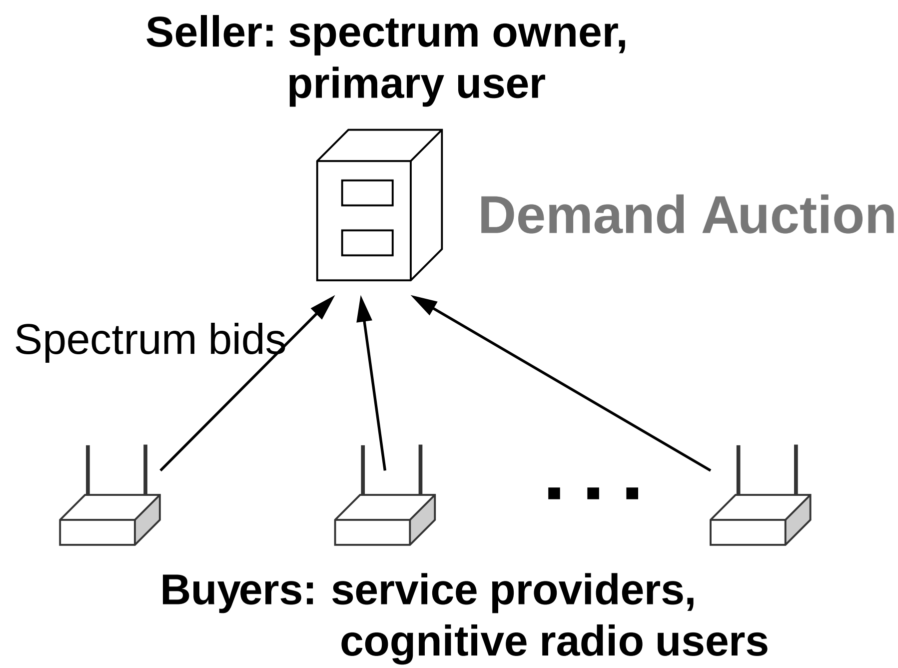
- Combinatorial auctions (CAs) are multi-item auctions in which bidders can place bids on combinations of items called packages or bundles, rather than just on individual item.
  - Pricing and bidding strategies
  - Winner determination problem
- Reverse CA game
  - The auctioneer first announces an initial price
  - The bidders submit bids corresponding to the current price
  - A bidder should determine a bid for every package he is interested in.
  - The auctioneer collects bids and finds the highest utility. If the highest utility is non-negative, the auctioneer is willing to sell the items, otherwise, he will not sell them.
  - Then the auctioneer reduces the price and the auction moves to the next round.
  - When stopping rules are satisfied, the auction is concluded.

Auction Types



Auction and Spectrum Management

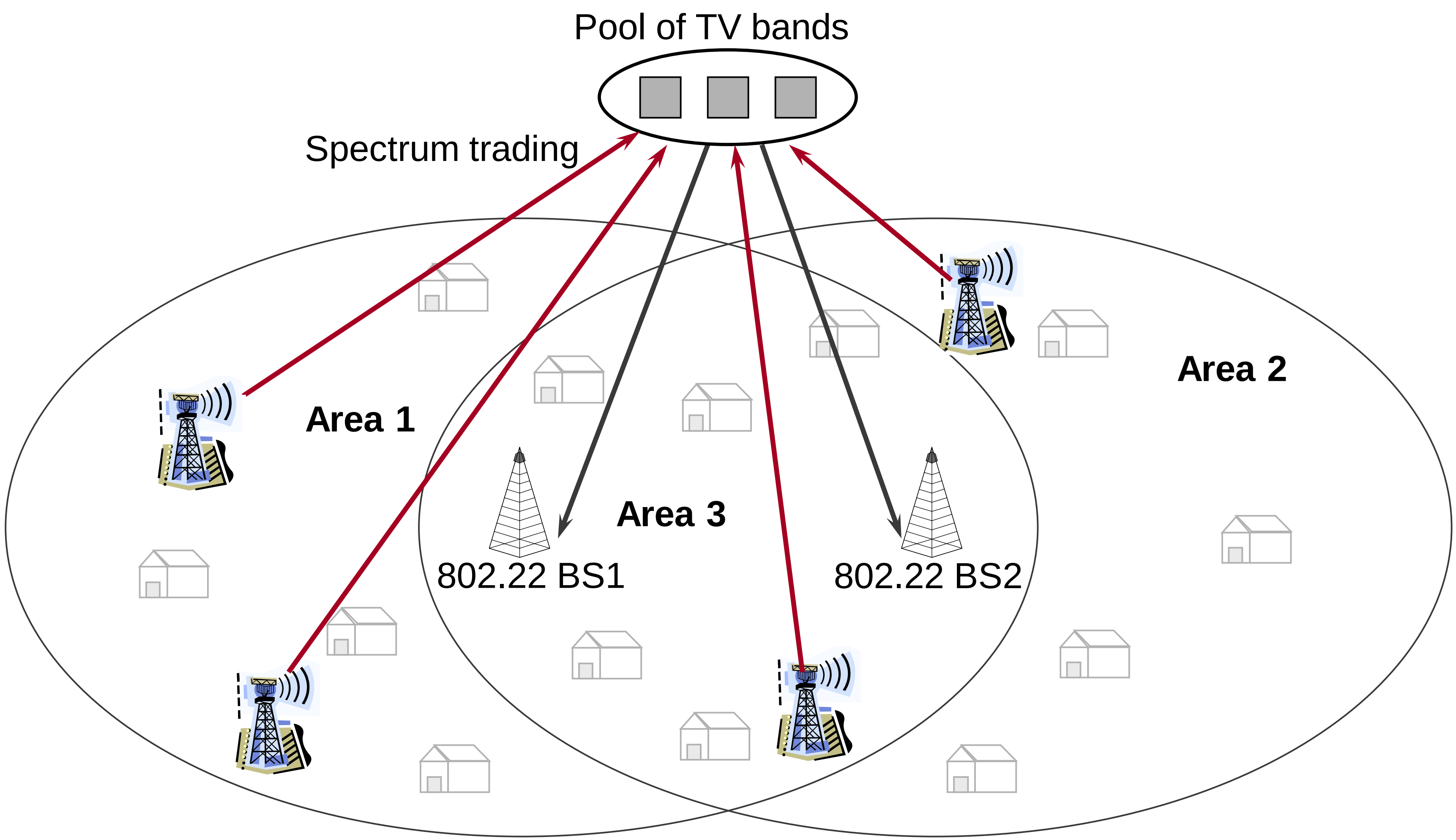
- Theory of auction can be applied to the problem of spectrum access and sharing (i.e., spectrum trading)



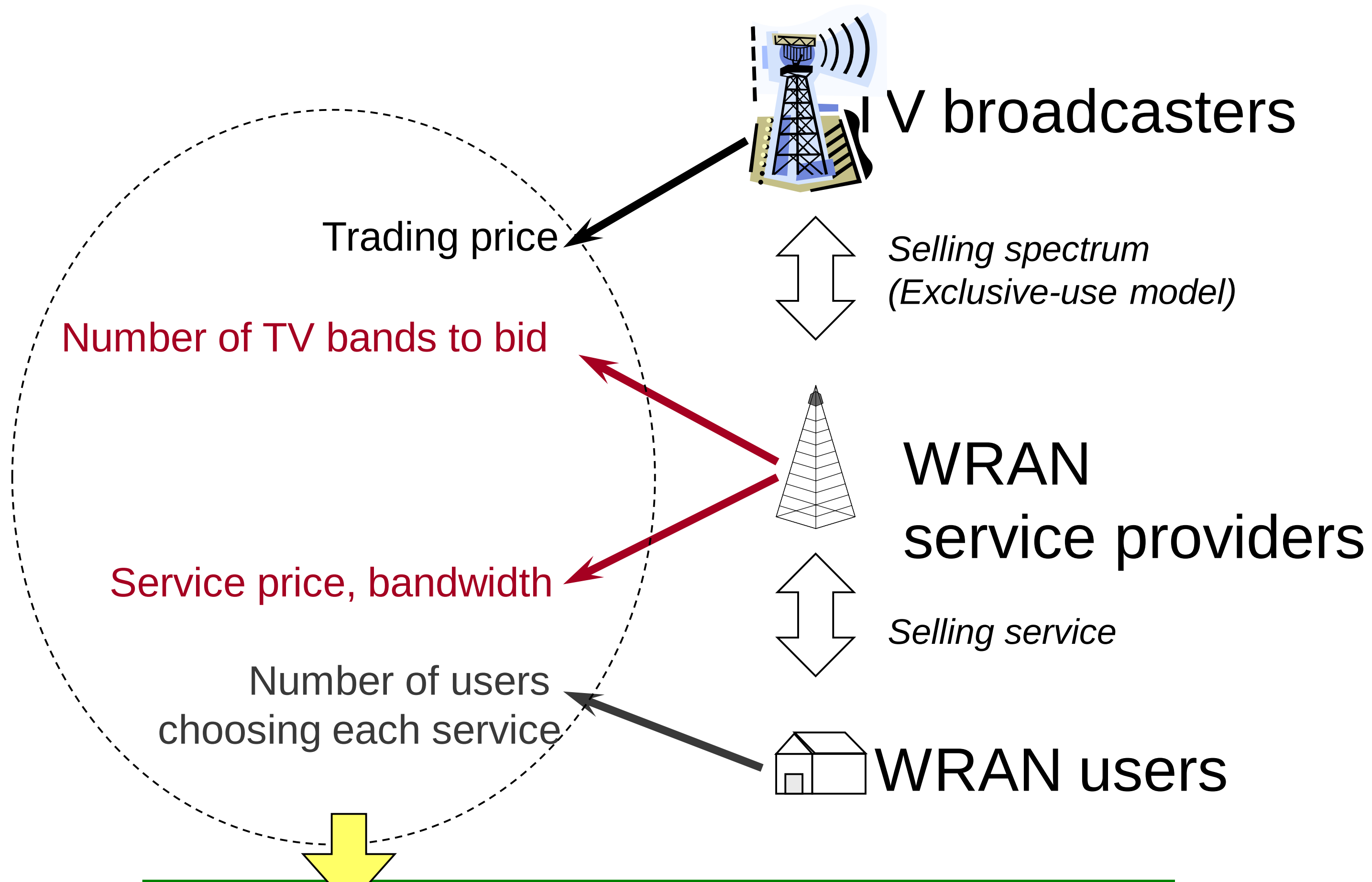
Auction and Spectrum Management

- Joint competitive spectrum bidding and service pricing in IEEE 802.22 networks [SM33]
  - Sellers: spectrum owner (i.e., TV broadcasters)
  - Buyers: IEEE 802.22 network service providers

System Model

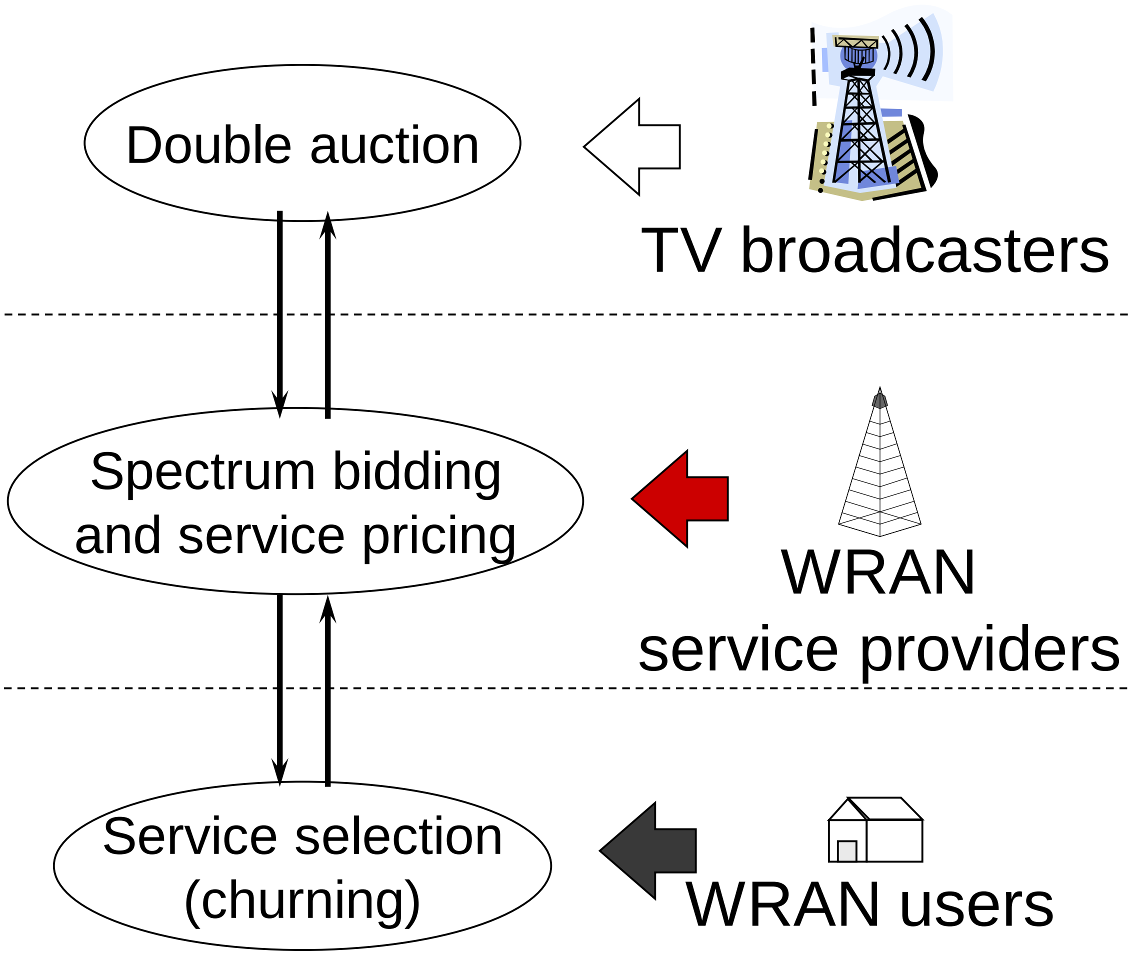


Problem Formulation

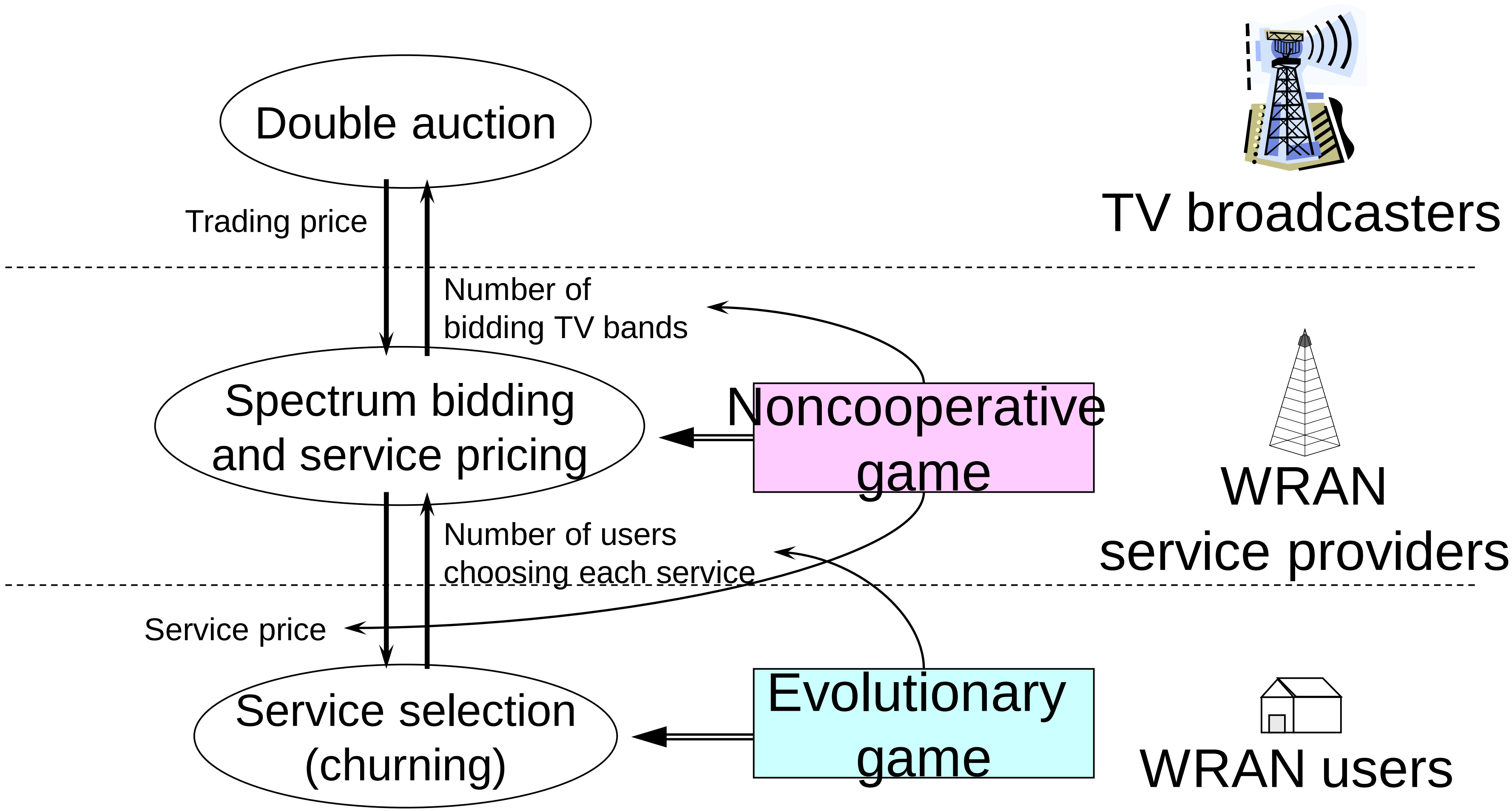


How to obtain the solutions?

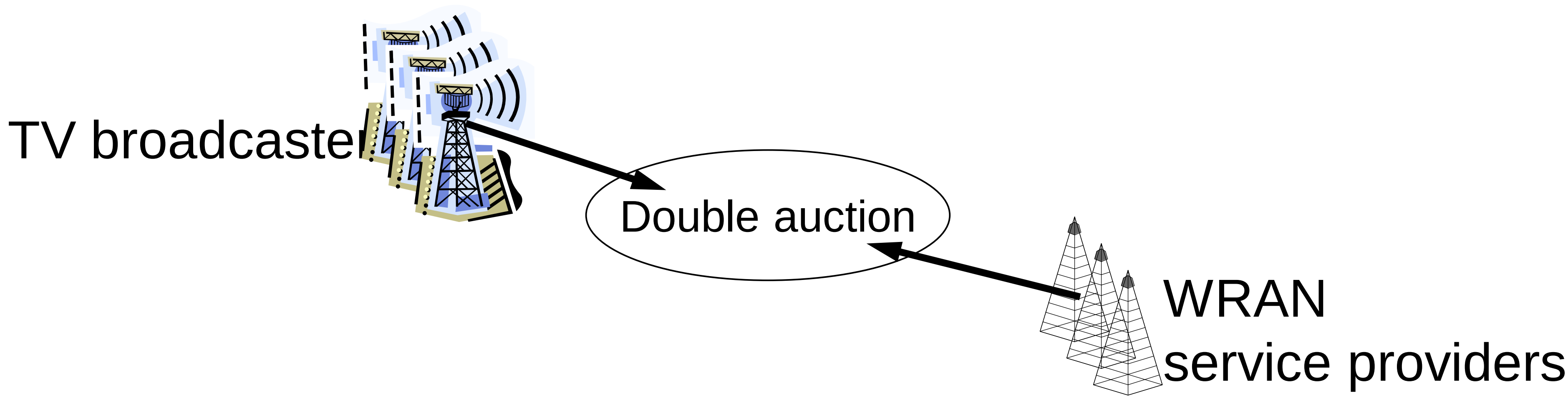
Problem Formulation



Game Model Formulations



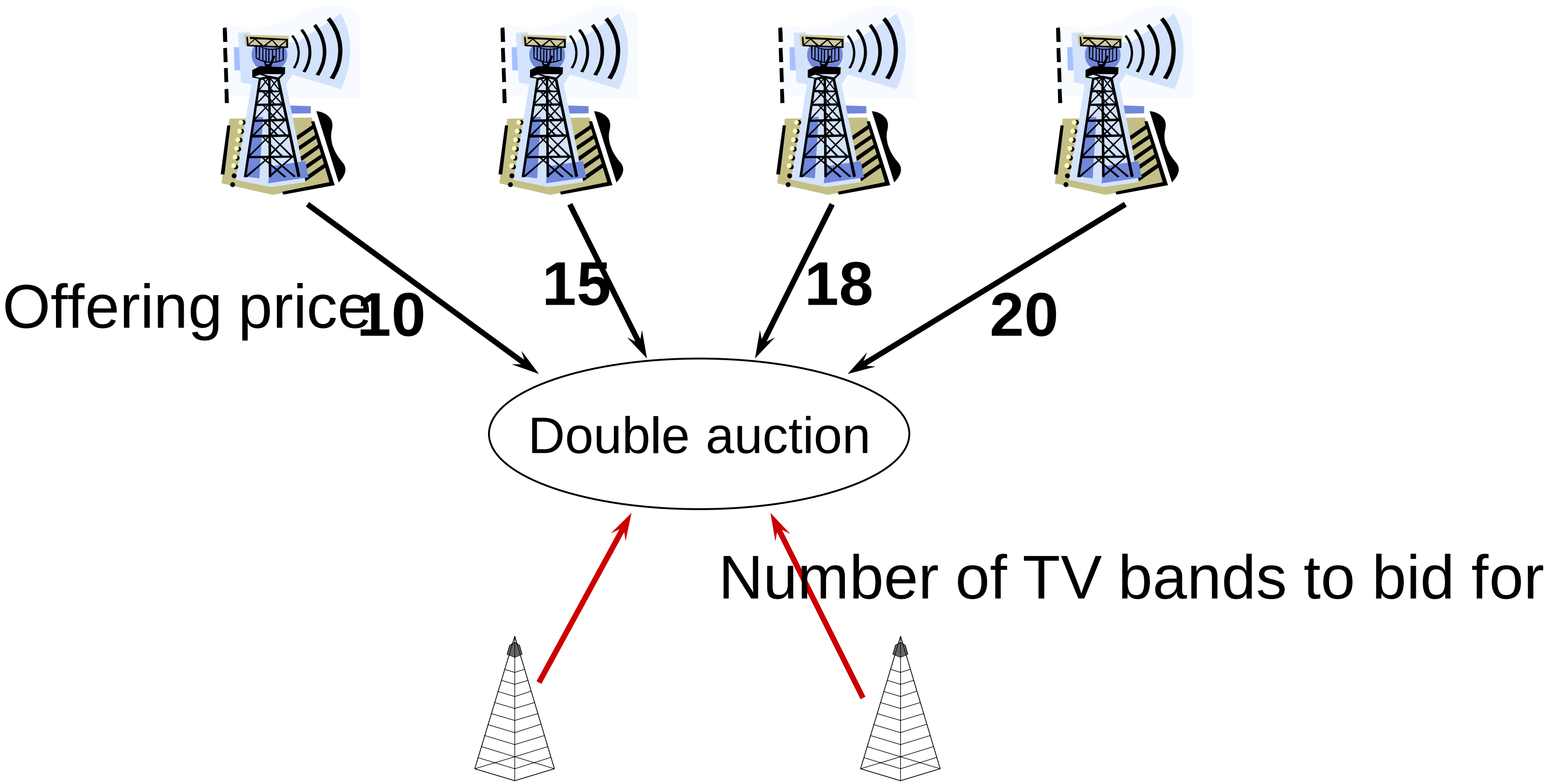
Double Auction



- Double auction is established among TV broadcasters (i.e., sellers) and WRAN service providers (i.e., buyers)
- Price of TV band is varied
- The market structure is for multiple-seller and multiple-buyers

Double Auction

- Example



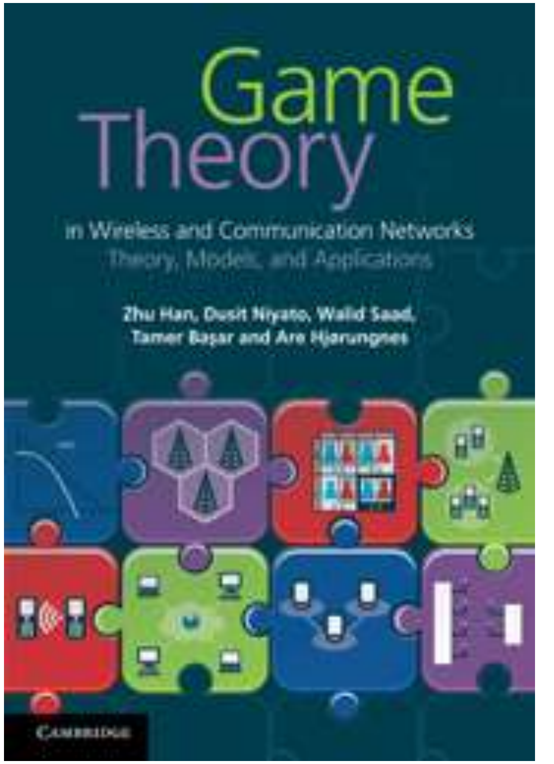
Summary

- Auctions take many forms but always satisfy two conditions:
  - They may be used to sell any item and so are *universal*, also
  - The outcome of the auction does not depend on the identity of the bidders; i.e., auctions are *anonymous*.
- Properties
  - efficiency of a given auction design
  - optimal and equilibrium bidding strategies
  - revenue comparison.
- Winner curse
  - winner will frequently have bid too much for the auctioned item.
- Many applications in wireless networking

Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 4
Evolutional Game

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of the Lecture Notes

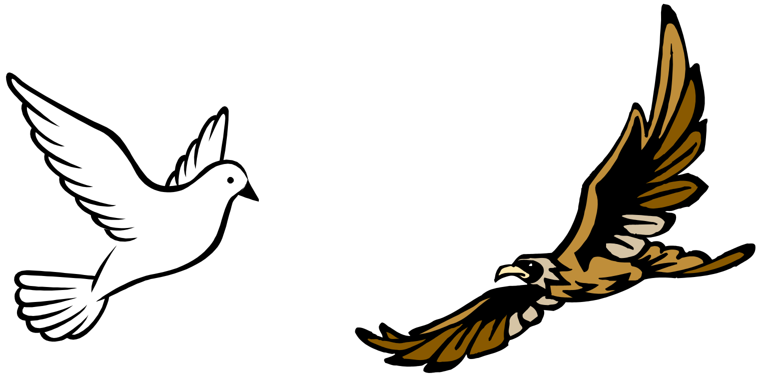
- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, book 2
- Contract Theory, Lecture 8, book 2
- Stochastic Game, Lecture 9, book 2
- Learning in Game, Lecture 10, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 11, book 2
- Mean Field Game, Lecture 12, book 2
- Zero Determinant Strategy, Lecture 13, book 2
- Network Economy, Lecture 14, book 2
- Game in Machine Learning, Lecture 15, book 2

Overview

- Basics of evolutionary
  - Equilibrium selection, bounded rationality, and dynamic behavior of players
- Two approaches in the evolutionary game framework
  - Static: evolutionary stable strategies (ESS)
  - Dynamic: replicator dynamics with evolutionary equilibrium
- Some of these applications have been discussed
  - Congestion control
  - Power control in CDMA,
  - Cooperative sensing in cognitive radio
  - Service provider selection (i.e., churning)

Overview of Evolutional Game

- Evolutionary game theory has been developed as a mathematical framework to study the interaction among rational **biological agents** in a population
- Agent adapts (i.e., evolves) the chosen strategy based on its fitness (i.e., payoff)
- Example, hawk (be aggressive) and dove (be mild)



Overview of Evolutional Game

- Evolutionary game theory has the following advantages over the traditional noncooperative game theory
  - The solution of the evolutionary game (i.e., evolutionary stable strategies (ESS) or evolutionary equilibrium) can serve as a **refinement** to the Nash equilibrium (e.g., Nash equilibrium is not necessarily efficient, there could be multiple Nash equilibria in a game, or the Nash equilibrium may not exist)
  - The **strong rationality** assumption is **not required** in evolutionary game as evolutionary game theory has been developed to model the behavior of biological agents
  - Evolutionary game is based on an evolutionary process, which is **dynamic** in nature which can model and capture the adaptation of agents to change their strategies and reach equilibrium over time

Evolution Process

- In an evolutionary game, the game is played repeatedly by agents who are selected from a large population
- Two major mechanisms of the evolutionary process and the evolutionary game are mutation and selection
  - Mutation is a mechanism of modifying the characteristics of an agent (e.g., genes of the individual or strategy of player), and agents with new characteristics are introduced into the population
  - The selection mechanism is then applied to retain the agents with high fitness while eliminating agents with low fitness
- In evolutionary game, mutation is described by the evolutionary stable strategies (ESS) from static system perspective
- Selection mechanism is described by the replicator dynamics from dynamic system perspective

Evolutionary Stable Strategies (ESS)

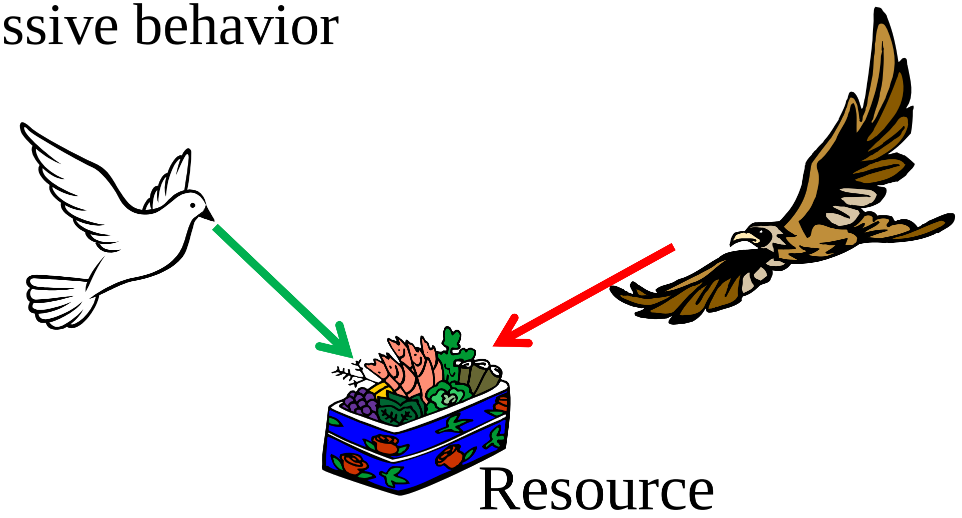
- ESS is the key concept in the evolutionary process in which a group of agents choosing one strategy will **not be replaced** by other agents choosing a different strategy when the mutation mechanism is applied
- Initial group of agents in a population chooses incumbent strategy *s*
- Small group of agents whose population share is ε choosing a different mutant strategy s'
- Strategy *s* is called evolutionary stable if

$$\boxed{u(s, \epsilon s' + (1-\epsilon)s) > u(s', \epsilon s' + (1-\epsilon)s),}$$

where $u(s, s')$ denote the payoff of strategy *s* given that the opponent chooses strategy *s*'

Example: Hawk-Dove Game

- There are two types of agents competing for a resource (i.e., food) of fixed value *V*
- Each agent chooses strategy from a set of two possibilities (i.e., hawk and dove)
  - Hawk is aggressive and will not stop fighting until it is injured or until the opponent retreats
  - Dove is mild behavior and always retreat instantly if the opponent initiates aggressive behavior



Example: Hawk-Dove Game

- There are 4 cases
  - 1) Both agents adopt hawk behavior (i.e., aggressive), the competition will result in both being equally injured with cost C
  - 2) One adopts hawk another adopts dove; dove immediately retreats and earns zero payoff, while the hawk captures the resource V
  - 3) When both adopt dove behavior, they will share the resource equally ($V/2$)
- Payoff matrix

	Hawk	Dove
Hawk	$1/2(V-C), 1/2(V-C)$	$V, 0$
Dove	$0, V$	$V/2, V/2$

- Almost all agents in the population adopt evolutionary stable strategy, no mutant (i.e., a small number of agents adopting a different strategy) can invade

Example: Hawk-Dove Game

- Illustration
  - Let $\varphi(s_1, s_2)$ denote the change in fitness for an agent adopting strategy s_1 against opponent adopting strategy s_2, and let $f(s)$ denote the total fitness of an agent adopting strategy s
  - Let f_0 denote the initial fitness, s denote the ESS, and s' denote the mutant strategy
  - The fitness of the agents adopting the different strategies can be express as follows:

$$
\begin{aligned}
f(s) &= f_0 + (1-\epsilon)\phi(s,s) + \epsilon\phi(s,s'), \\
f(s') &= f_0 + (1-\epsilon)\phi(s',s) + \epsilon\phi(s',s'),
\end{aligned}
$$

  - Where ε is proportion of the population for the mutant strategy s'

Example: Hawk-Dove Game

- Illustration (Cont.)
  - For ESS, the fitness of the agent adopting strategy *s* must be larger than that of those members of the population choosing strategy *s*' (i.e., $f(s) > f(s')$)
  - If ε approaches zero, it is required that either of these conditions holds, i.e.,

$$\begin{aligned} \phi(s,s) &> \phi(s',s), \\ \phi(s,s) &= \phi(s',s) \text{ and } \phi(s,s') > \phi(s',s') \end{aligned}$$

- For Hawk-Dove game, the <u>dove is not ESS</u> since a pure population of doves can be invaded by a hawk mutant

- If resource V is larger than the cost of both agents behaving aggressively (i.e., $V > C$), then the <u>hawk is ESS</u> as there is value in both agents competing for a resource even though they would be hurt

- Otherwise, there is no ESS in this game

Replicator Dynamics

- Population can be divided into multiple groups, and each group adopts a different pure strategy
- Replicator dynamics can model the evolution of the group size over time (unlike ESS, in replicator dynamics agents will play only pure strategies)
- The proportion or fraction of agents using pure strategy s (i.e., population share) is denoted by $x_s(t)$ whose vector is $\mathbf{x}(t)$
- Let payoff of an agent using strategy s given the population state $\mathbf{x}$ be denoted by $u(s, \mathbf{x})$
- Average payoff of the population, which is the payoff of an agent selected randomly from a population, is given by

$$\boxed{\overline{u}(\mathbf{x}) = \sum_{s \in \mathcal{S}} x_s u(s, \mathbf{x})}$$

Replicator Dynamics

- The reproduction rate of each agent (i.e., the rate at which the agent switches from one strategy to another) depends on the payoff (agents will switch to strategy that leads to higher payoff)
- Group size of agents ensuring higher payoff will grow over time because the agents having low payoff will switch their strategies
- Dynamics (time derivative) of the population share can be expressed as follows:

$$\boxed{\dot{x}_s = x_s\left(u(s,\mathbf{x}) - \overline{u}(\mathbf{x})\right)}$$

- Evolutionary equilibrium can be determined at $\dot{x}_s = 0$

  where actions of the population choosing different strategies cease to change

Replicator Dynamics

- It is important to analyze the stability of the replicator dynamics to determine the evolutionary equilibrium
- Evolutionary equilibrium can be stable (i.e., equilibrium is robust to the local perturbation) in the following two cases:
  - 1) Given the initial point of replicator dynamics sufficiently close to the evolutionary equilibrium, the solution path of replicator dynamics will remain arbitrarily close to the equilibrium (Lyapunov stability)
  - 2) Given the initial point of replicator dynamics close to the evolutionary equilibrium, the solution path of replicator dynamics converges to the equilibrium (asymptotic stability)
- Two main approaches to prove the stability of evolutionary equilibrium are based on the *Lyapunov function* and the *eigenvalue* of the corresponding matrix

Example: Prisoner's Dilemma

- Two agents choose a strategy of cooperate or defect

	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

  where $T > R > P > S$

- x_C and x_D denote the proportions of the population adopting cooperate and defect strategies, respectively
- Average fitness of agents adopting these two strategies are denoted by u_C and u_D, respectively
- Average fitness of the entire population is $\overline{u}$ obtained from

$$
\begin{aligned}
u_C &= u_0 + x_C\phi(C, C) + x_D\phi(C, D), \\
u_D &= u_0 + x_C\phi(D, C) + x_D\phi(D, D), \\
\overline{u} &= x_C u_C + x_D u_D,
\end{aligned}
$$

Change in fitness

Example: Prisoner's Dilemma

- The future proportion of the population adopting the strategies (i.e., $\hat{x}_C$ and $\hat{x}_D$) depends on the current proportion

Cooperate	*Defect*
$\hat{x}_C = \frac{x_C u_C}{\overline{u}}$, $\hat{x}_C - x_C = \frac{x_C (u_C - \overline{u})}{\overline{u}}$,	$\hat{x}_D = \frac{x_D u_D}{\overline{u}}$, $\hat{x}_D - x_D = \frac{x_D (u_D - \overline{u})}{\overline{u}}$

- Consider small time interval, the differential equations (replicator dynamics) are

$$\boxed{\frac{dx_C}{dt} = \dot{x}_C \approx \frac{x_C (u_C - \overline{u})}{\overline{u}}} \qquad \boxed{\frac{dx_D}{dt} = \dot{x}_D \approx \frac{x_D (u_D - \overline{u})}{\overline{u}}}$$

Example: Prisoner's Dilemma

- For the prisoner's dilemma case, we have $u_C = u_0 + x_C R + x_D S$ and $u_D = u_0 + x_C T + x_D P$
- Since $T > R$ and $P > S$, it is clear that $u_D > u_C$, and

$$\boxed{\frac{u_D - \overline{u}}{\overline{u}} > 0 \text{ and } \frac{u_C - \overline{u}}{\overline{u}} < 0}$$

- Therefore, as time increases, the proportion of the population adopting the cooperate strategy will approach zero (i.e., becomes extinct)
- From replicator dynamics, **defect** strategy constitutes the evolutionary equilibrium
- Also, it can be proven that defect strategy is the ESS of the prisoner's dilemma game

Overview

- Basics of evolutionary
  - Equilibrium selection, bounded rationality, and dynamic behavior of players
- Two approaches in the evolutionary game framework
  - Static: evolutionary stable strategies (ESS)
  - Dynamic: replicator dynamics with evolutionary equilibrium
- Some of these applications have been discussed
  - Congestion control
  - Power control in CDMA,
  - Cooperative sensing in cognitive radio
  - Service provider selection (i.e., churning)

Applications of Evolutionary Game

Congestion control

- The competition among two types of behaviors (i.e., aggressive and peaceful) in wireless nodes to access the channel using a certain protocol can be modeled as an evolutionary game
- Congestion control is (transport layer) to avoid performance degradation by the ongoing users by limiting transmission rate
- The transmission rate (i.e., of TCP) can be adjusted by changing the congestion window size (i.e., the maximum number of packets to be transmitted)
- The speed-of-transmission rate to be increased and decreased defines the aggressiveness of the protocol

Applications of Evolutionary Game

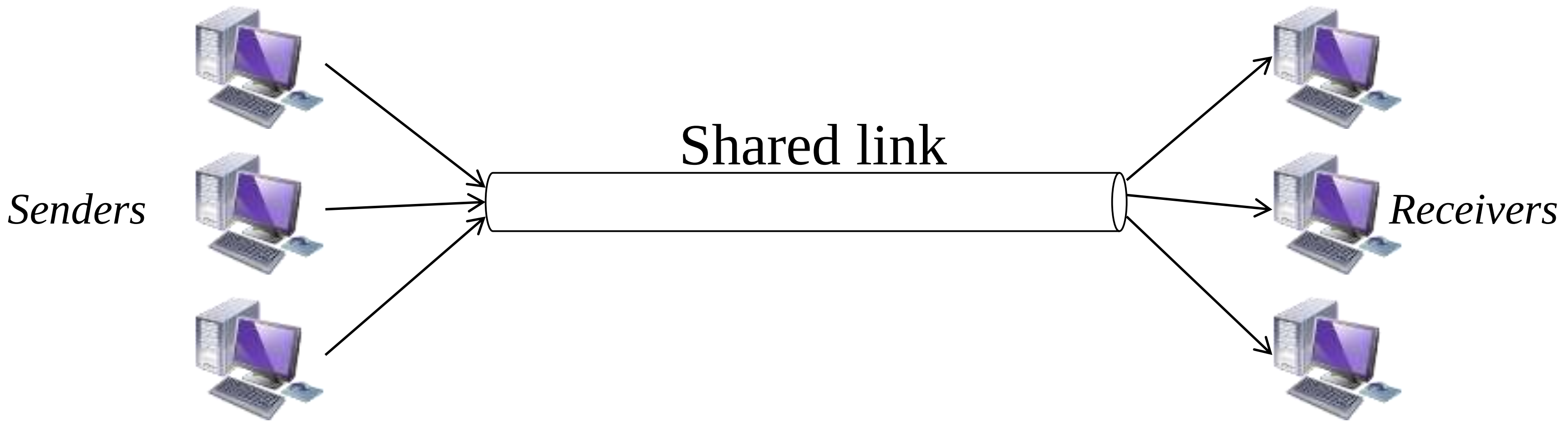
Congestion control

- TCP protocol with the additive increase multiplicative decrease (AIMD) mechanism can control this aggressiveness through the parameters determining the increase and decrease
- If the transmitted packet is successful, the window size will linearly increase by α packets for every round trip time
- Otherwise, the window size will decrease by β proportional to the current size

Applications of Evolutionary Game

Congestion control

- Multiple flows share the same link, competitive situation arises



- It is found that the aggressive strategy of all flows (i.e., large values of α and β) becomes the Nash equilibrium, and the performance will degrade significantly due to the congestion

Applications of Evolutionary Game

Congestion control – Static game

- Analysis of the TCP protocol in a wireless environment is performed in which the evolutionary game model (similar to the Hawk and Dove game)
- There are two populations (i.e., groups) of flows with TCP
- The flow from population *i* is characterized by parameters α_i and β_i, which are the increase and decrease rates, respectively
- Strategy *s* of flow is to be aggressive (i.e., hawk or *H*) to be peaceful (i.e., dove or *D*)
- The parameters associated with these strategies are given as

$$(\alpha_i, \beta_i) \in \{(\alpha_\mathsf{H}, \beta_\mathsf{H}), (\alpha_\mathsf{D}, \beta_\mathsf{D})\} \text{ for } \alpha_\mathsf{H} \geq \alpha_\mathsf{D}$$

Applications of Evolutionary Game

Congestion control – Static game

- The packet loss occurs when the total transmission rate of all flows reaches the capacity C- i.e., $x_1 r_1 + x_2 r_2 = C$, where x_i is the proportion of population choosing aggressive behavior
- The payoff of flow in population i is defined as follows:

$$u_i = \tau_i - \omega L,$$

  where τ_i is the average throughput, L is the loss rate, and ω is the weight for the loss
- Throughput of flow from population i can be obtained from

$$\tau_i = \frac{1+\beta_i}{2} \frac{\alpha_i \overline{\beta}_j}{\alpha_i \overline{\beta}_j + \alpha_j \overline{\beta}_i} C \qquad \overline{\beta}_i = 1 - \beta_i$$

Applications of Evolutionary Game

Congestion control – Static game

- The average throughput and loss rate can be defined as functions of strategies of two populations i.e., $\tau_i(s_i, s_j)$ and $L(s_i, s_j)$
- It is shown that $\tau_i(H, H) = \tau_i(D, D)$
- When the loss rate is considered, it increases as the flow becomes more aggressive, i.e., larger values of α_i and β_i
- Therefore, it can be shown that $u_i(H, H) < u_i(D, D)$ and $u_i(D, H) < u_i(D, D)$
- Game becomes a Hawk and Dove model whose solution is ESS
- Briefly, it is found that the application that is loss-sensitive will tend to use a less aggressive strategy at ESS

Applications of Evolutionary Game

Congestion control – Dynamic game

- Dynamics of strategy selection by the flows in two populations can also be analyzed using the replicator dynamics

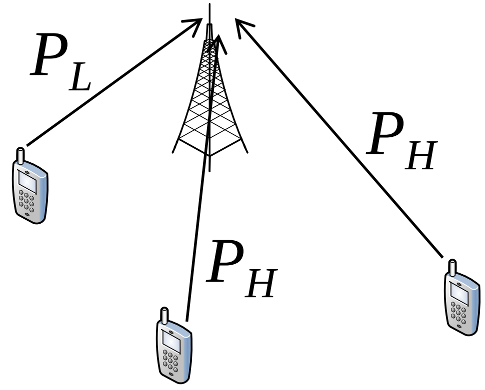
$$\boxed{\dot{x}_s(t) = \frac{\mathrm{d}x_s(t)}{\mathrm{d}t} = x_s(t)K\left(u(s, \mathbf{x}(t)) - \sum_{s'} x_{s'}(t)u(s', \mathbf{x}(t))\right)}$$

x_s is the proportion of the population choosing strategy s and $x_s(t)$ is a vector of x_s at time t; $u(s, x(t))$ is the payoff of using strategy s, and K is a speed constant (positive)

Applications of Evolutionary Game

Evolutionary Game for WCDMA Access

- Evolutionary game is formulated for the WCDMA system
- The number of interfering nodes is random, which depends on the geographical location of the mobile nodes
- Mobile nodes have two strategies to use high and low power levels, which correspond to the transmit power P_H and P_L, respectively



Applications of Evolutionary Game

Evolutionary Game for WCDMA Access

- Signal-to-interference-plus-noise ratio (SINR) with distance r between transmitter and receiver of node i is given by

$$\gamma_i(P_i, x, r) = \begin{cases} \frac{gP_i/r_0^\alpha}{\sigma+\beta I(x)}, & \text{if } r \leq r_0, \\ \frac{gP_i/r^\alpha}{\sigma+\beta I(x)}, & \text{if } r > r_0, \end{cases}$$

  - P_i is the strategy of node i (i.e., P_H or P_L)
  - x is the proportion of the population choosing P_H
  - g is channel gain, r_0 is the radius-of-reception circle of receiver
  - α is the attenuation order with value between 3 and 6, σ is the noise power, and β is the inverse of processing gain
  - $I(x)$ is total interference from all nodes to the receiver of node i

Applications of Evolutionary Game

Evolutionary Game for WCDMA Access

- Payoff of node i is as follows:

$$\boxed{u_i(P_i, x) = \int_0^R \log(1 + \gamma_i(P_i, x, r))\zeta(r)\mathrm{d}r - w_\mathrm{p} P_i,}$$

  - R is the transmission range, and w_p is the cost weight due to adopting transmit power P_i (e.g., energy consumption)
  - $\zeta(r)$ is the probability density function given the density of receiver

Applications of Evolutionary Game

Evolutionary Game for WCDMA Access

- Based on this evolutionary game formulation, the sufficient condition for existence and uniqueness of the ESS in WCDMA access is established
- Dynamics of the evolutionary game formulation of WCDMA access can be established based on replicator dynamics

$$\boxed{\dot{x}(t) = x(t)(1 - x(t))\left(\mathscr{H}(x(t)) - w_{\mathrm{p}}(P_{\mathrm{H}} - P_{\mathrm{L}})\right).}$$

$$\mathscr{H}(x) = \log\left(\frac{1 + \gamma_i(P_{\mathrm{H}}, x, r)}{1 + \gamma_i(P_{\mathrm{L}}, x, r)}\right)\zeta(r).$$

This function is continuous and strictly monotonic, which is required for the proof of stability based on sufficient condition

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- In a cognitive radio network, unlicensed users (i.e., secondary users) performs spectrum sensing to detect licensed users (i.e., primary users) before opportunistically access the spectrum
- It is based on sampling the signal with hypotheses that a primary user is present or absent denoted by H_1 and H_0, respectively
- Multiple secondary users can cooperate and share the sensing results to reduce the sensing time while maintaining the detection and false-alarm probabilities at the target levels
- However, there will be the secondary users who contribute or deny to contribute in cooperative spectrum sensing because they are rational

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- Secondary users denying to participate in cooperative spectrum sensing will have more time for data transmission
- However, if none of the secondary users performs cooperative sensing, the throughput will be low because the detection probability is low and false-alarm probability is high
- This conflict situation can be analyzed using the evolutionary game framework

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- The evolutionary game is defined as follows
  - Players are the secondary users (i.e., totally N players)
  - Strategies are to contribute or deny, which are denoted by C and D, respectively
  - The payoff is the throughput of the secondary user defined as follows:

$$\boxed{u_{C,i} = P_{H_0}\left(1 - \frac{T_{\mathrm{sense}}}{|\mathcal{C}|\,T}\right)(1 - P_{\mathrm{fal}}(\mathcal{C}))R_i,}$$

  P_{H0} is the probability of the spectrum to be idle (i.e., a primary user is absent)

  C is a set of contributing secondary users

  $P_{fal}(C)$ is the false-alarm probability given a set of contributing secondary users C, and R_i is the transmission rate of user i

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- For denying secondary user *j*, the payoff function is

$$\boxed{u_{D,j} = P_{H_0}\left(1 - P_{\text{fal}}(\mathcal{C})\right) R_j .}$$

  – Since the denying secondary users do not need to spend time for sensing, their throughput is large

- Replicator dynamics is

$$\boxed{\dot{x}_i = \left(u_{C,i}(\mathbf{x}_{-i}) - \overline{u}_i(\mathbf{x})\right) x_i ,}$$

x_i denote the probability of secondary user *i* selecting a contributing strategy

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- For homogeneous case, all secondary users are taken to be identical (i.e., the same detection and false alarm probabilities, and the same transmission rate), average payoffs are

$$\overline{u}_{\mathrm{C}} = \sum_{j=0}^{F-1} \binom{F-1}{j} x^j (1-x)^{F-1-j} u_{\mathrm{C}}(j+1), \quad \textit{Cooperate}$$

$$\overline{u}_{\mathrm{D}} = \sum_{j=0}^{F-1} \binom{F-1}{j} x^j (1-x)^{F-1-j} u_{\mathrm{D}}(j). \quad \textit{Deny}$$

F is the number of channels

Applications of Evolutionary Game

Cooperative Sensing in Cognitive Radio

- Replicator dynamics can be modified to

$$\boxed{\dot{x} = x(1-x)\left(\overline{u}_{\mathrm{C}} - \overline{u}_{\mathrm{D}}\right).}$$

- Also, evolutionary stable strategies (ESS) can be obtained as the solution of x^* for $\dot{x} = 0$ by solving the following equation

$$\boxed{\frac{T_{\mathrm{sense}}}{T}(1-x^*)^F + Fx^*(1-x^*)^{F-1} - \frac{T_{\mathrm{sense}}}{T} = 0.}$$

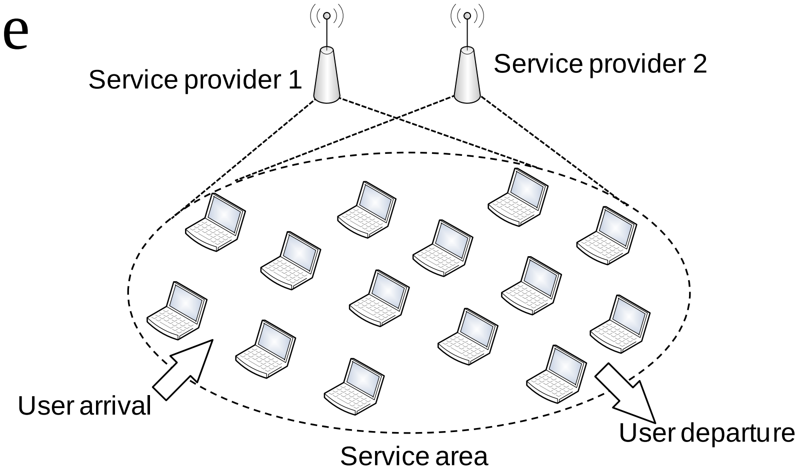
T_{sense} is the time interval for sensing

T is the length of time slot

Applications of Evolutionary Game

Mobile User Churning Behavior

- Churning of mobile users is common since mobile users have freedom to choose the best wireless service
- Churning behavior of wireless service users is analyzed using the theory of evolutionary games
- WLAN hotspot is considered where a wireless user can choose among different IEEE 802.11-based WLAN access points based on the performances and/or price



Applications of Evolutionary Game

Mobile User Churning Behavior

- Mobile users' behavior
  - User tends to choose and churn to the wireless service provider that returns a higher payoff
  - Due to the lack of information about the performance obtained from different service providers and/or inadequate information about the decisions of other users, a user has to gradually learn and change decision on choosing a particular wireless service
  - A user can make a wrong decision to choose a wireless service provider that provides a lower payoff randomly with a small probability
  - An individual user does not have any intention to influence the decisions of other users in the service area

Applications of Evolutionary Game

Mobile User Churning Behavior

- The payoff of a user choosing wireless service provider s

$$\boxed{u(s) = \mathscr{U}(\tau_s(n_s)) - p_s,}$$

$\mathscr{U}(\tau_s)$ is concave utility (logarithmic) function of throughput τ_s

p_s is a price charged by service provider s to a user

- Throughput is obtained from (standard IEEE 802.11 formula)

$$\tau_s(n_s) = \frac{P_{\text{succ}}\overline{\text{Pack}}}{\overline{\Phi} + P_{\text{succ}} T_{\text{succ}} + (1 - P_{\text{succ}}) T_{\text{coll}}},$$

Applications of Evolutionary Game

Mobile User Churning Behavior

- Stochastic Dynamic Evolutionary Game Formulation
  - Connections are initiated at an average rate of λ
  - Holding time is exponentially distributed with mean $1/\mu$
  - Demand function, the effective connection arrival rate is

$$\boxed{\tilde{\lambda} = \lambda \exp\left(-\left(\frac{\sum_{s=1}^{S} p_s/S}{p_0} - 1\right)^2\right)}$$

S is total number of service providers

p_0 is normal price

Applications of Evolutionary Game

Mobile User Churning Behavior

- Stochastic dynamic evolutionary game can be modeled as a continuous-time Markov chain
- State space of this Markov chain can be described as follows:

$$\Delta = \{(N_1, \dots, N_s, \dots, N_S) | 0 \leq N_s \leq N\}$$

  N_s is the number of users selecting service provider s

  N is the total number of users in a service area

- The transition rate can be derived given following events
  - Connection arrival and departure
  - Rational and irrational churning

Applications of Evolutionary Game

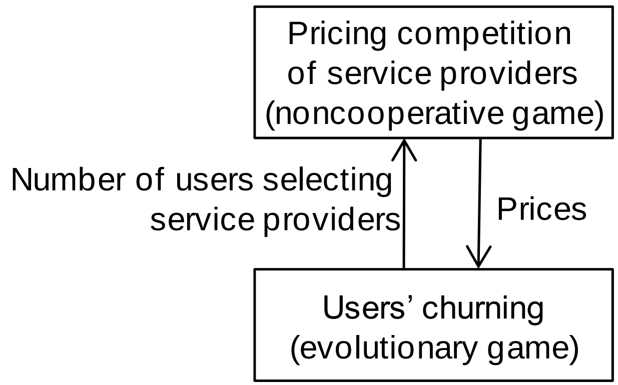
Mobile User Churning Behavior

- Rational churning happens with rate $n_s\left(u(s') - u(s)\right)$ from service provider *s* to *s*'
  - User changes to service provider yielding higher payoff *u*(.)
- Then, steady-state probability of Markov chain can be obtained which determines the probability of having n_s users for service provider *s*

Applications of Evolutionary Game

Mobile User Churning Behavior

- Given the model of churning behavior, the competitive pricing can be analyzed



- Revenue earned by service provider *s* given price p_s is

$$\boxed{\mathscr{R}_s(p_s, \mathbf{p}_{-s}) = \overline{n}_s(p_s, \mathbf{p}_{-s})p_s,}$$

$\overline{n}_s(p_s, \mathbf{p}_{-s})$ is average number of users choosing service provider *s* (obtained from evolutionary game model)

Applications of Evolutionary Game

Mobile User Churning Behavior

- Solution of this price competition among the service providers is the Nash equilibrium, for which the condition is

$$\boxed{\mathscr{R}_s(p_s^*, \mathbf{p}_{-s}^*) \geq \mathscr{R}_s(p_s, \mathbf{p}_{-s}^*),}$$

- <u>Cooperative Pricing:</u> all wireless service providers agree (i.e., collude) to choose the price so that their revenue is maximized

$$\boxed{\mathbf{p}^* = \arg\max_{\mathbf{p}} \sum_{s=1}^{S} \mathscr{R}_s(\mathbf{p}),}$$

Summary

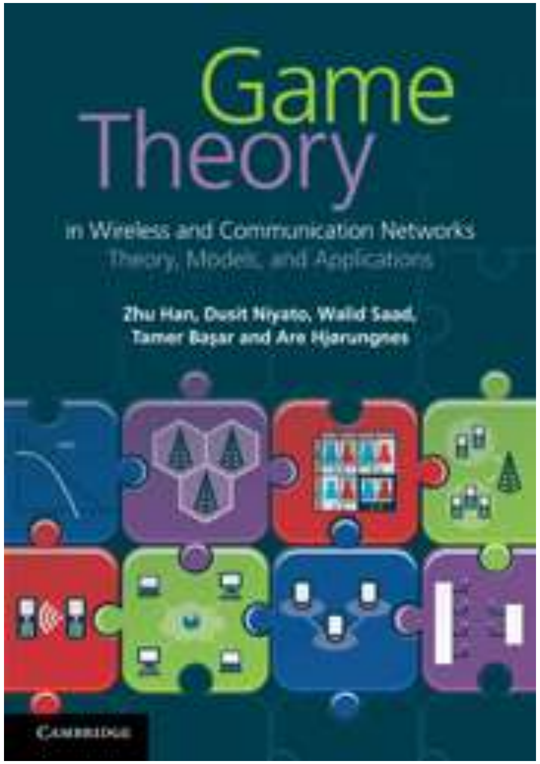
- Basics of evolutionary games are presented and its advantages over the classical noncooperative game are discussed
  - Equilibrium selection, bounded rationality, and dynamic behavior of players
- Two approaches in the evolutionary game framework
  - Static: evolutionary stable strategies (ESS)
  - Dynamic: replicator dynamics with evolutionary equilibrium
- Some of these applications have been discussed
  - Congestion control
  - Power control in CDMA,
  - Cooperative sensing in cognitive radio
  - Service provider selection (i.e., churning)

Game Theory in Wireless and Communication Networks: Theory, Models, and Applications

Lecture 1

Noncooperative Games

Zhu Han, Dusit Niyato, Walid Saad,
Tamer Basar, and Are Hjorungnes



Overview of the Lecture Notes

- Introduction to Game Theory: Lecture 1, book 1
- Non-cooperative Games: Lecture 1, Chapter 3, book 1
- Bayesian Games: Lecture 2, Chapter 4, book 1
- Differential Games: Lecture 3, Chapter 5, book 1
- Evolutionary Games: Lecture 4, Chapter 6, book 1
- Cooperative Games: Lecture 5, Chapter 7, book 1
- Auction Theory: Lecture 6, Chapter 8, book 1
- Matching Game: Lecture 7, book 2
- Contract Theory, Lecture 8, book 2
- Stochastic Game, Lecture 9, book 2
- Learning in Game, Lecture 10, book 2
- Equilibrium Programming with Equilibrium Constraint, Lecture 11, book 2
- Mean Field Game, Lecture 12, book 2
- Zero Determinant Strategy, Lecture 13, book 2
- Network Economy, Lecture 14, book 2
- Game in Machine Learning, Lecture 15, book 2

History of Game Theory

- **John von Neuman** (1903-1957) co-authored, Theory of Games and Economic Behavior, with Oskar Morgenstern in 1940s, establishing game theory as a field
- **John Nash** (1928-) developed a key concept of game theory (Nash equilibrium) which initiated many subsequent results and studies
- Since 1970s, game-theoretic methods have come to dominate microeconomic theory and other fields
- Nobel prizes
  - Nobel prize in Economic Sciences 1994 awarded to **Nash**, **Harsanyi** (Bayesian games) and **Selten** (Subgame perfect equilibrium)
  - 2005, **Auman** and **Schelling** got the Nobel prize for having enhanced our understanding of cooperation and conflict through game theory
  - 2007, **Leonid Hurwicz**, **Eric Maskin** and **Roger Myerson** won Nobel Prize for having laid the foundations of mechanism design theory

John von Neumann
(December 28,1903 – February 8,1957)

John Forbes Nash, Jr.
(born June 13, 1928)
Winner of Nobel Prize in Economics (1994)

Introduction

- ***Game Theory:*** Mathematical models and techniques developed in economics to analyze interactive decision processes, predict the outcomes of interactions, and identify optimal strategies
- Game theory techniques were adopted to solve many protocol design issues (e.g., resource allocation, power control, cooperation enforcement) in wireless networks
- Difference to control: against other players as well as nature
- Fundamental component of game theory is the notion of a game
  - A game is described by a set of rational *players*, the *strategies* associated with the players, and the *payoffs* for the players. A rational player has his own interest, and therefore, will act by choosing an available strategy to achieve his interest.
  - A player is assumed to be able to evaluate exactly or probabilistically, the outcome or payoff (usually measured by the utility) of the game which **depends not only on his action but also on other players' actions**.

Examples: Rich Game Theoretical Approaches

- Non-cooperative Static Games (Mandayam and Goodman, 2001)
  - Sports: Zero sum game, Boxing: Example of equilibrium
  - Virginia tech
- Repeated Games (MAD: Nobel prize, 2005)
  - Threat of punishment by repeated games (playing multiple times)
  - Tit-for-Tat (Infocom, 2003)
- Dynamic Games (Basar's book)
  - ODE for state, Optimization utility over time, HJB and dynamic programming
  - Evolutional game (Hossain and Dusit's work)
- Stochastic Games (Altman's work)
- Cooperative Games
  - Nash Bargaining Solution
  - Coalitional Game

Auction Theory

- Book of Myerson (Nobel Prize 2007), J. Huang, H. Zheng, X. Li

Contract Theoretical: Asymmetry Information

- Noble Prize 2014: Adverse Selection and Moral Hazard
- ***Moral Hazard of my PhD student***

What my parents thinks I do

What my advisor thinks I do

- What I actually do

When advisor presents

When advisor on travel

Matching Game: GS algorithm

2012 Nobel Prize in Economic Science.

Adam

Geeta, Heiki, Irina, Fran

Bob

Irina, Fran, Heiki, Geeta

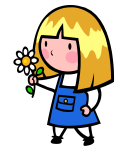
Carl

Geeta, Fran, Heiki, Irina

Davi
d

Irina, Heiki, Geeta, Fran

We reach a stable marriage!

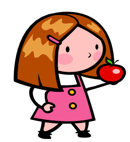


Fran

Geeta

Carl > Adam

Heiki



Irina

David > Bob

Overview of Non-cooperative Game Theory

- Basics
- Games in Strategic (Normal) Form
  - Dominating Strategy
  - Nash Equilibrium
  - Mixed Strategy
  - Static Continuous-Kernel Games
- Dynamic Non-cooperative Games
  - Games in Extensive Form
  - Repeated Games
  - Stochastic Games
- Special Games
  - Potential Games
  - Stackelberg Games
  - Correlated Equilibrium
  - Supermodular Games
  - Wardrop Games
- Summary

Basics

- Involves a number of players having totally or partially conflicting interests on the outcome of a decision process
- ***Static Game:*** Players take their actions only once, independently of each other
- ***Dynamic Game:*** Players have some information about each others' choices, can act more than once
- ***Strategic (Normal) Form:*** One of the most popular representations

Games in Strategic (Normal) Form

- A game in strategic (normal) form is represented by three elements
  - A finite set of players, N
  - The set of available strategies for player i, S_i
  - The utility (payoff) function for player i, u_i
- $\mathbf{s} = (s_i , \mathbf{s}_{-i})$ is the strategy profile, where s_i € S_i is the strategy of player i, and $\mathbf{s}_{-i}$ the vector of strategies of all players except i
- Note that one user's utility is a function of both this user's and others' strategies
- ***Complete Information Game:*** If all elements of the game are common knowledge
- ***Incomplete Information Game:*** The players may not know the identities of all other players, their payoffs or their strategies

Matrix Form

- Game represented by a **matrix**
- Row represents a strategy for the first (row) player
- Column represents the strategies of the second (column) player
- Matrix entry, (x, y) where x is the payoff of the first player and y is the payoff of the second player
- Example: The Prisoner's Dilemma
  - Two suspects in a crime held for interrogation in separate cells
  - If they both stay quiet, each will be convicted with a minor offence and will spend **1 year** in prison
  - If one and only one of them finks, he will be freed and used as a witness against the other who will spend **4 years** in prison
  - If both of them fink, each will spend **3 years** in prison

Example: The Prisoner's Dilemma

- **Rational Players:** The prisoners
- **Strategies:** Stay Quiet (Q) or Fink (F)
- **Solution:** What is the Nash equilibrium of the game?

	P2 Quiet	P2 Fink
P1 Quiet	1,1	4,0
P1 Fink	0,4	3,3

Nash Equilibrium (1)

- ***Dominant Strategy:*** A player's best strategy, i.e., a strategy that yields the highest utility for the player regardless of what strategies the other players choose
- ***Nash Equilibrium:*** A strategy profile **s*** with the property that no player *i* can do better by choosing a strategy different from **s***, given that the strategies of other players $j \neq i$ remain fixed
- In other words, for each player *i* with payoff function u_i ,

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*), \forall s_i \in S_i$$

- No player can improve its payoff by **unilaterally** changing its strategy, given that the other players' strategies remain fixed

Nash Equilibrium (2)

- Does the Nash equilibrium always exist?
- Is it efficient?
- ***Pareto Optimality:*** A measure of efficiency
  - A payoff vector **x** is Pareto optimal if there does **not** exist any payoff vector **y** such that

    $$\mathbf{y} \geq \mathbf{x}$$

    with at least one strict inequality for an element y_i
- In some references a strategy profile **s** that achieves a Pareto optimal payoff distribution can sometimes be referred to as a **Pareto optimal strategy**

The Price of Anarchy

- ***Centralized System:*** Given a global knowledge of the parameters, one seeks to find the social optimum
  - Efficient but often unfair
- ***Decentralized System:*** When there is non-cooperation and competition among the players, one operating point of interest is the Nash equilibrium
  - Inefficient but stable from the players' perspective
- ***The Price of Anarchy (PoA):*** The ratio of the cost (or utility) function at social optimum with respect to the equilibrium case
  - Measures the price of not having a central coordination in the system
  - PoA is, loosely, a measure of the loss incurred by having a distributed system!

Example: The Prisoner's Dilemma

- The Price of Anarchy is 1/3

	P2 Quiet	P2 Fink
P1 Quiet	1,1	4,0
P1 Fink	0,4	3,3

Pareto optimal (recall we're minimizing)

Nash equilibrium

Example: Battle of Sexes

- Multiple Nash equilibriums

	Opera	Football
Opera	2,3	0,0
Football	0,0	3,2

Nash equilibrium

Nash equilibrium

Pure vs. Mixed Strategies

- So far we assumed that the players make **deterministic choices** from their strategy spaces
- ***Pure Strategies:*** If a player i selects a strategy out of its strategy set S_i, in a deterministic manner (with probability 1)
- ***Mixed Strategies:*** Players select a **probability distribution** over their set of pure strategies
- **Nash 1950**
  - Every finite non-cooperative game in strategic form has a **mixed strategy Nash equilibrium**

Mixed Nash Equilibrium

- Define σ_i as a probability mass function over S_i, the set of actions of player i
- When working with mixed strategies, each player i aims to maximize its **expected payoff**

$$u_i(\sigma) = \sum_{s \in S} \left(\prod_{j=1}^{I} \sigma_j(s_j) \right) u_i(s)$$

- A mixed strategy profile $\boldsymbol{\sigma}$* is a **Mixed strategy Nash equilibrium**, if for each player i, we have

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*) \qquad \textit{for all } \sigma_i \in \Sigma_i.$$

Example: Battle of Sexes

	Opera	Football
Opera	2,3	0,0
Football	0,0	3,2

- Husband picks Opera with probability *p* , wife picks Opera with probability *q*
- Expected payoff for husband picking Opera: 2q
- Expected payoff for husband picking Football: 3(1-q)
- At mixed NE, the expected payoff at a strategy is equal to that at another strategy (otherwise, one would use a pure NE)
- Mixed NE -> Husband: (2/5,3/5) Wife: (3/5,2/5)
- Expected payoffs (6/5,6/5)

Algorithms for Finding the NE

- For a general N-player game, finding the set of NEs is not possible in polynomial time!
    - *Unless the game has a certain structure*
- Some existing algorithms
    - Fictitious play (based on empirical probabilities)
    - Iterative algorithms (can converge for certain classes of games)
    - Best response algorithms
        - *Popular in some games (continuous kernel games for example)*
    - Useful reference
        - *D. Fundenberg and D. Levine, The theory of learning in games, the MIT press, 1998.*

Static Continuous-Kernel Game

- Action (strategy) sets have uncountably many elements
  - For example, strategies are intervals
- Similar to the definitions before
  - **Best response**

Definition 9 *The* best response function $b_i(\mathbf{s}_{-i})$ *of a player i to the profile of strategies $\mathbf{s}_{-i}$ is a set of strategies for that player such that*

$$b_i(\mathbf{s}_{-i}) = \{s_i \in \mathcal{S}_i \mid u_i(s_i, \mathbf{s}_{-i}) \geq u_i(s_i', \mathbf{s}_{-i}),\ \forall s_i' \in \mathcal{S}_i\}. \tag{3.5}$$

  - **Nash equilibrium**

Proposition 1 *A strategy profile $\mathbf{s}^* \in \mathcal{S}$ is a* Nash equilibrium *of a noncooperative game if and only if every player's strategy is a* best response *to the other players' strategies, that is:*

$$s_i^* \in b_i(\mathbf{s}_{-i}^*) \text{ for every player } i. \tag{3.6}$$

Example: Power Control Game

- Interference channel SINR

$$\gamma_i(\mathbf{p}) = \frac{p_i h_i}{\sigma^2 + \sum_{j \neq i} p_j h_j},$$

Definition 10 *A function* $\mathbf{g} : \mathcal{S} \to \mathbb{R}_+^N$ *is said to be standard if it has the following properties:*

- *Monotonicity:* $\forall \mathbf{s}, \mathbf{s}' \in \mathcal{S},\ \mathbf{s} \leq \mathbf{s}' \Rightarrow \mathbf{g}(\mathbf{s}) \leq \mathbf{g}(\mathbf{s}')$ *(component-wise).*
- *Scalability:* $\forall \alpha > 0, \mathbf{s} \in \mathcal{S}, \mathbf{g}(\alpha \mathbf{s}) \leq \alpha \mathbf{g}(\mathbf{s})$.

Theorem 2 ***(Yates 1995)*** *If the best response functions of a noncooperative game* G*, are standard functions for all players, i.e.,* $\forall i \in \mathcal{N}$*, then, the game has a unique Nash equilibrium in pure strategies.*

Guidance to Design Utility Function

- Existence and uniqueness of Nash equilibrium depend on the utility function design
  - Shows the existence of pure strategy Nash equilibria

Theorem 3 *(Debreu - Fan - Glicksberg 1952) Given a noncooperative game in strategic form* $(\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (u_i)_{i \in \mathcal{N}})$, *if* $\forall i \in \mathcal{N}$, *every strategy set* $\mathcal{S}_i$ *is compact and convex,* $u_i(s_i, \mathbf{s}_{-i})$ *is a continuous function in the profile of strategies* $\mathbf{s} \in \mathcal{S}$ *and quasi-concave in* s_i, *then the game has at least one pure strategy Nash equilibrium.*

  - Shows the existence of a **unique** pure strategy Nash equilibrium

Theorem 4 *(Rosen 1965) Consider a strategic game* G, *where* $\forall i \in \mathcal{N}$, *every strategy set* $\mathcal{S}_i$ *is compact and convex,* $u_i(s_i, \mathbf{s}_{-i})$ *is a continuous function in the profile of strategies* $\mathbf{s} \in \mathcal{S}$ *and concave in* s_i. *Let* $\mathbf{r} = (r_1, \ldots, r_N)$ *be an arbitrary vector of fixed positive parameters. If the Diagonal Strict Concavity (DSC) property holds true, i.e.,*

$$\exists \mathbf{r} > 0 : (\mathbf{s} - \mathbf{s}')(g(\mathbf{s}, \mathbf{r}) - g(\mathbf{s}', \mathbf{r})) > 0, \ \forall \mathbf{s}, \mathbf{s}' \in \mathcal{S}, \ \mathbf{s} \neq \mathbf{s}', \tag{3.12}$$

with $g(\mathbf{s}, \mathbf{r}) \triangleq [r_1 \frac{\partial u_1(s_1, \mathbf{s}_{-1})}{\partial s_1}, \ldots, r_N \frac{\partial u_N(s_N, \mathbf{s}_{-N})}{\partial s_N}]^T$, *then, the game has a* unique pure strategy Nash equilibrium.

Overview of Non-Cooperative Game Theory

- Basics
- Games in Strategic (Normal) Form
  - Dominating Strategy
  - Nash Equilibrium
  - Mixed Strategy
  - Static Continuous-Kernel Games
- Dynamic Non-cooperative Games
  - Games in Extensive Form
  - Repeated Games
  - Stochastic Games
- Special Games
  - Potential Games
  - Stackelberg Games
  - Correlated Equilibrium
  - Supermodular Games
  - Wardrop Games
- Summary

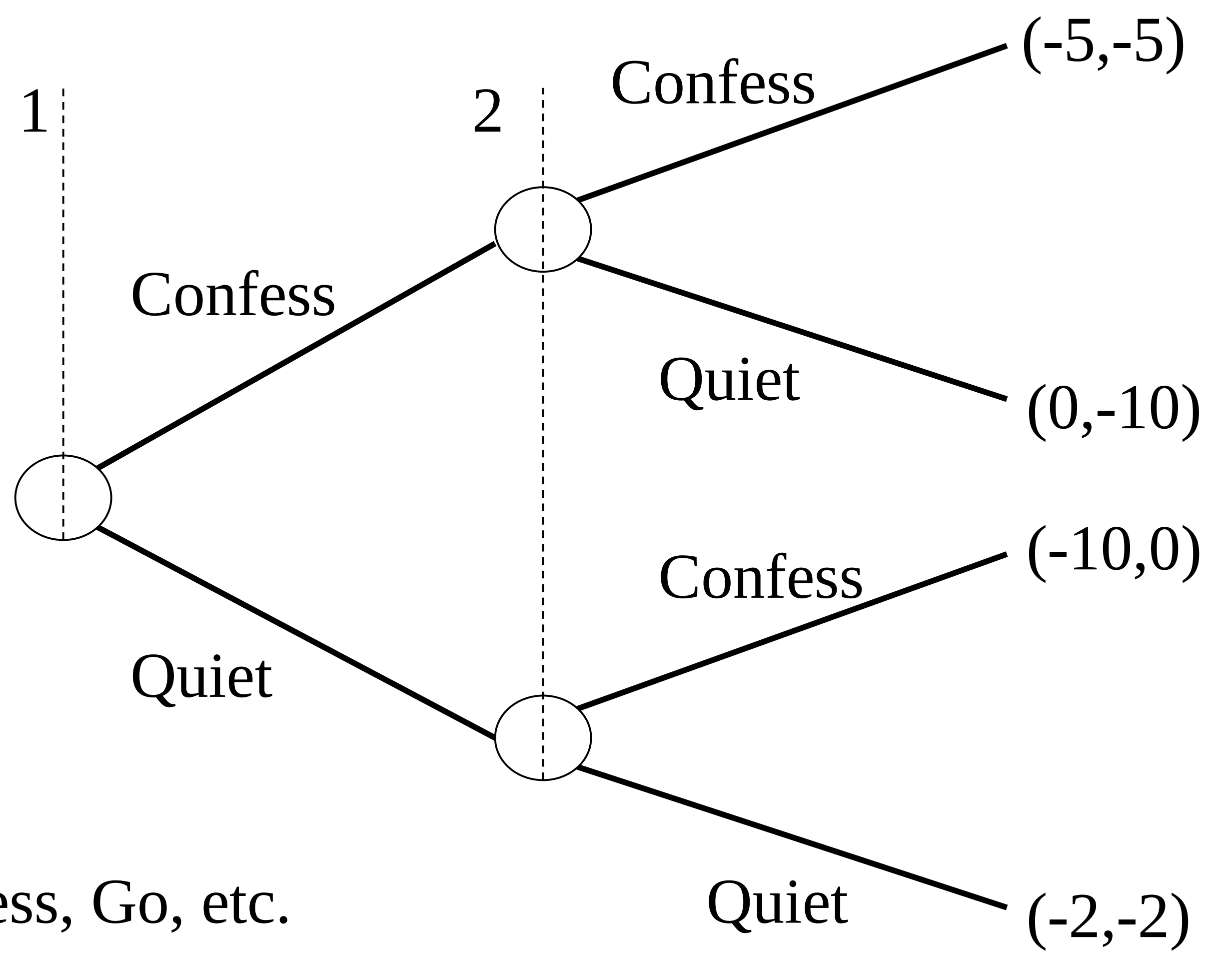
Games in Extensive Form (1)

- In dynamic games, the **notion of time and information is important**
  - The strategic form cannot capture this notion
  - We need a new game form to visualize a game
- ***Sequential Games:*** A major class of dynamic games in which players take their decisions in a certain pre-defined order
  - ***Perfect Information:*** Only one player moves at a time and each player knows every action of the players that moved before it at every point of the game
  - ***Imperfect Information:*** Some of the players do not know all the previous choices of the other players
- In extensive form, a game is represented with a game tree
  - Provides not only a representation of the players, payoffs, and their actions, but also a representation of the order of moves (or sequence) and the information sets

Games in Extensive Form (2)

- ***Game Tree:***
  - Nodes: Points at which players can take actions. Each node is labeled so as to identify who is making the decision.
    - **Initial (or root) node:** The **first decision** to be made by one of the players.
    - **Terminal node:** Every set of vertices from the first node through the tree arrives at this node, representing an end to the game. Each terminal node is labeled with the **payoffs** earned by each player.
  - Edges: Nodes are connected by edges, which represent the **moves** that may be taken by a player present at a certain node.
  - Information Sets: A number of nodes enclosed by dotted lines which represent the **information available** for the player at the time of play.
    - **Perfect information game:** Every information set contains exactly one node since each player knows exactly all the information.
    - **Imperfect information game:** There exists at least one information set containing more than one node.
  - Stage: One level of the tree.
  - History: Sequence of actions that were taken up to the considered stage.

Example: The Prisoner's Dilemma



Chess, Go, etc.

Games in Extensive Form (3)

- Normal form gives only the minimum amount of information necessary to describe a game
- The extensive form gives additional details about the game
  - The timing of the decisions to be made and the amount of information available to each player when each decision has to be made
- For every extensive form game, there is one and only one corresponding normal form game
- For every normal form game, there are, in general, several corresponding extensive form games
- Every finite extensive form game of perfect information has a pure strategy Nash equilibrium

Subgame

- A subgame of a dynamic noncooperative game consists of a **single node in the extensive form** representation of the game
  - The game tree, and all of its successors down to the terminal nodes
- The information sets and payoffs of a subgame are inherited from the original game
- Moreover, the strategies of the players are restricted to the history of actions in the subgame

Definition 14 *A strategy profile* $\mathbf{s} \in S$ *is a* subgame perfect equilibrium *if players' strategies (restricted to a subgame) constitute a Nash equilibrium in every subgame of the original game.*

Subgame Perfect Equilibrium

- Backward induction is a useful method to find equilibria in a dynamic game (in extensive form) with perfect information.

Definition 15 *The* one-stage deviation principle *requires that: There must not exist any information set, in which a player i can gain by deviating from its subgame perfect equilibrium strategy (at this information set) while its strategy at other information sets as well as the strategies of the other players are fixed.*

In other words, a strategy profile $\mathbf{s}^*$ is a subgame perfect equilibrium if for each player $i \in \mathcal{N}$ and at each information set where player i moves if we:

- Fix the other players strategies as $\mathbf{s}^*$.
- Fix player i's moves at other information sets as in $\mathbf{s}^*$.

Then, player i cannot improve its payoff (at the information set) by deviating from s_i at the information set only. Note that for games with perfect information the above definitions reduce to backward induction.

Finding Subgame Perfect Equilibrium

- Steps:
  1. Pick a subgame that does not contain any other subgame.
  2. Compute a Nash equilibrium of this subgame.
  3. Assign the payoff vector associated with this equilibrium to the starting node, and eliminate the subgame.
  4. Iterate this procedure until a move is assigned at every contingency, when there remains no subgame to eliminate.
- Nash equilibrium is not necessarily a subgame perfect equilibrium.

Example: Sequential Multiple Access Game

- A basic technique in CSMA/CA protocol
- The two devices *p1* and *p2* are not perfectly synchronized
  - *p1 decides to transmit or not*
  - *p2 observes p1 before making his own move*
- The strategy of *p1* is to *Transmit (T)* or to be *Quiet (Q)*
- How many pure Nash equilibria do we have?
- H. W. Kuhn, "Extensive Games and the problem of information", Contributions to the Theory of Games II, 1953.

p_1	p_2	payoff
T	T	(-c,-c)
T	Q	(1-c,0)
Q	T	(0,1-c)
Q	Q	(0,0)

- ***Theorem (Kuhn, 1953):*** Every finite extensive form game of perfect information has a pure strategy Nash equilibrium.

Backward Induction for Sequential Multiple Access Game

- How do we solve the game?
- If player *p2* plays the strategy *T* then the best response of player *p1* is to play *Q*.
- However, *T* is not the best strategy of player *p2* if player *p1* chooses *T*.
- We can eliminate some possibilities by backward induction.
- Player ***p2* knows that he has the last move**.
- **Given all the best moves of *p2*, player *p1* calculates his best moves** as well.
- It turns out that the backward induction solution is the historical move (T) then (Q).

p_1	p_2	payoff
T	T	(-c,-c)
T	Q	(1-c,0)
Q	T	(0,1-c)
Q	Q	(0,0)

Example: Without Backward Induction

- Imperfect information; backward induction **not possible**
  - Consider the two subgames: The subgame that starts after Player 1 plays D and the subgame which is the game itself
  - Consider the first subgame and find NE
  - Clearly only one NE: **(L,U)**, Payoff: **(3, 2)**
  - Game tree reduced to the second one
  - Clearly *player 1* would choose **D**
  - Subgame perfect equilibrium:

    Player 1 picks **D** at stage 1

    L at stage 2

    Player 2 always picks **U**

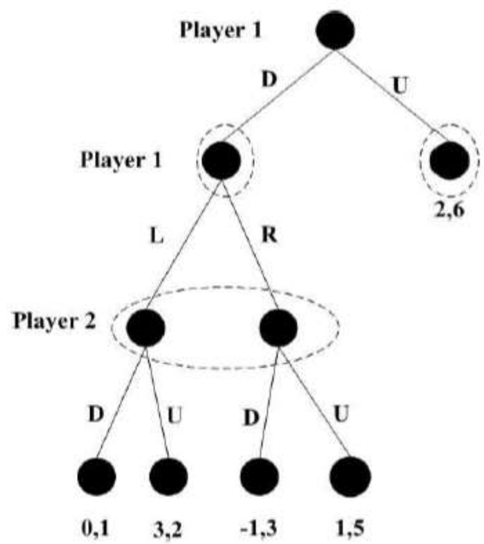
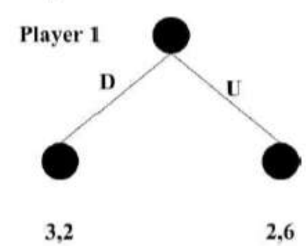


Figure 3.3: A Dynamic Game In Extensive Form.



Repeated Game Basics

- Static noncooperative strategic game that is repeated over time
- **Average utility** (power, in our case) over time
  - Discounting factor, β

$$u_i = (1-\beta)\sum_{t=1}^{\infty} \beta^{t-1} u_i(t)$$

- ***Folk Theorem:***
  - A feasible outcome can be obtained through a repeated game
  - In the long run, the players, although acting noncooperatively, might choose a cooperative behavior so as to obtain a payoff that is better than the minmax value.
- Enforcing cooperation by punishment: Trigger Strategies
  - ***Tit-for-Tat:*** Player responds by defecting, once the opponent defects
  - ***Cartel Maintenance:*** Provide enough threat to greedy players
- **Approach:** Cooperate initially, detect the outcome of the game
  - If better than a threshold, play cooperation the next time. Else, play non-cooperation for a period T, and then cooperate.

Stochastic Game Basics

- A repeated game with **stochastic (probabilistic) transitions** between the different states of the game.
- A dynamic game composed of a number of stages, and where, at the beginning of each stage the game, is in some state.
- In this state, the players select their actions and each player receives a payoff that depends on the current state and the chosen actions.
- The game then moves to a new random state whose distribution depends on the previous state and the actions chosen by the players.
- The procedure is repeated at the new state and the game continues for a finite or infinite number of stages.
- The total payoff to a player is often taken to be the discounted sum of the stage payoffs (similar to the discounted sum of repeated games) or the limit inferior of the averages of the stage payoffs.
- Notice that 1-state stochastic game is equal to (infinitely) repeated game, and 1-agent stochastic game is equal to Markov Decision Process (MDP).
- Partial observed MDP is a widely used model for wireless networking.
- Bellman equation

Overview of Non-Cooperative Game Theory

- Basics
- Games in Strategic (Normal) Form
  - Dominating Strategy
  - Nash Equilibrium
  - Mixed Strategy
  - Static Continuous-Kernel Games
- Dynamic Non-cooperative Games
  - Games in Extensive Form
  - Repeated Games
  - Stochastic Games
- Special Games
  - Potential Games
  - Stackelberg Games
  - Correlated Equilibrium
  - Supermodular Games
  - Wardrop Games
- Summary

Potential Games (1)

- A special class of noncooperative games having a special structure
- Non-zero sum games in which the determination of a Nash equilibrium can be equivalently posed as a maximization of a single function, called the potential function
- The potential function is a useful tool to analyze **equilibrium properties** of games
  - Since the objectives and goals of all players are mapped into one function
- Potential games are characterized by their simplicity and the existence of a Nash equilibrium solution
- Often, potential games are useful when dealing with continuous-kernel games

Potential Games (2)

- Formally,

Definition 18 *A noncooperative strategic game* $(\mathcal{N}, (\mathcal{S}_i)_{i\in\mathcal{N}}, (u_i)_{i\in\mathcal{N}})$ *is an exact (cardinal) potential game if there exists an exact potential function* $\Phi : \mathcal{S} \to \mathbb{R}$ *such that* $\forall i \in \mathcal{N}$

$$\Phi(x, \mathbf{s}_{-i}) - \Phi(z, \mathbf{s}_{-i}) = u_i(x, \mathbf{s}_{-i}) - u_i(z, \mathbf{s}_{-i}), \ \forall x, z \in \mathcal{S}_i, \forall \mathbf{s} \in \mathcal{S}. \qquad (3.28)$$

A game is a general (ordinal) potential game if there is an ordinal potential function $\Phi : \mathcal{S} \to \mathbb{R}$ *such that*

$$\mathrm{sgn}[\Phi(x, \mathbf{s}_{-i}) - \Phi(z, \mathbf{s}_{-i})] = sgn[u_i(x, \mathbf{s}_{-i}) - u_i(z, \mathbf{s}_{-i})], \ \forall x, z \in \mathcal{S}_i, \forall \mathbf{s} \in \mathcal{S}, \qquad (3.29)$$

where sgn *denotes the sign function.*

- In *exact potential games*, the difference in individual utilities has the same value as the difference in potential function values
  - Achieved by each player when changing its strategy *unilaterally*
- In *ordinal potential games*, only the signs of the differences have to be the same

Potential Games (3)

Corollary 1 *Every finite potential game (exact or ordinal) has at least one pure strategy Nash equilibrium.*

Corollary 2 *For infinite potential games (with finite number of players), a pure strategy Nash equilibrium exists if: (i) $\mathcal{S}_i$ are compact strategy sets and (ii) the potential function Φ is upper semi-continuous on $\mathcal{S}$.*

Theorem 9 *Given a strategic game where the strategy sets $\mathcal{S}_i$, $\forall i \in \mathcal{N}$ are intervals of real number and assuming the utilities are twice continuously differentiable, then this game is a potential game if and only if*

$$\frac{\partial^2 (u_i - u_j)}{\partial s_i \partial s_j} = 0 \ \forall i \in \mathcal{N}, j \in \mathcal{N}. \qquad (3.30)$$

- The Nash equilibrium is **unique**
  - When the strategy set S is compact and convex and the potential function is a continuously differentiable function on the interior of S and strictly concave on S

Example: Single Cell CDMA Network

- Consider a single-cell CDMA network, M users
- SINR

$$\gamma_i(\boldsymbol{p}) = \frac{p_i h_i}{n_0 + \sum_{j \neq i} p_j h_j}$$

- Optimization

$$\begin{array}{ll} \text{minimize} & p_i \\ \text{subject to} & f_i(\gamma_i(\boldsymbol{p})) \geq \gamma_i^{thresh} \\ \text{variables} & p_i \in [0, P_i^{\max}]. \end{array}$$

a non-cooperative game $G = \left[\mathcal{M}, \mathcal{A}, \{\log(P_i^{\max} - p_i)\}_{i \in \mathcal{M}}\right]$, with the *coupled* action set is

$$\mathcal{A} = \left\{\boldsymbol{p} : f_i(\gamma_i(\boldsymbol{p})) \geq \gamma_i^{thresh}, p_i \in [0, P_i^{\max}], \forall i \in \mathcal{M}\right\}.$$

Furthermore, the game G admits a potential function

$$Z(\boldsymbol{p}) = \sum_{i \in \mathcal{M}} \log(P_i^{\max} - p_i).$$

We can then maximize function $Z(\boldsymbol{p})$ over set $\mathcal{A}$, and the corresponding maximizer(s) will be the NE(s) of game G, and thus the optimal solution(s) of Problem (1.4) for all users.

Stackelberg Games (1)

- **Hierarchy** among the players exists
  - The player that imposes its own strategy upon others is called the **leader**
  - The other players who react to the leader's strategy are called **followers**

Definition 20 *In a two-person finite game with Player* 1 *as the leader, a strategy* $s_1^* \in \mathcal{S}_1$ *is called a* Stackelberg equilibrium strategy *for the leader, if*

$$\min_{s_2 \in \mathcal{R}_2(s_1^*)} u_1(s_1^*, s_2) = \max_{s_1 \in \mathcal{S}_1} \min_{s_2 \in \mathcal{R}_2(s_1)} u_1(s_1, s_2) \triangleq u_1^*. \tag{3.35}$$

The quantity u_1^* *is the* Stackelberg utility *of the leader. The same definition applies for the case where player* 2 *is the leader by simply swapping the subscripts* 1 *and* 2.

- Every two-person finite game admits a Stackelberg strategy for the leader
- Whenever the follower has a single optimal response for every strategy of the leader, then the leader can, at the Stackelberg solution, perform **at least as good as at the Nash equilibrium**

Stackelberg Games (2)

- Stackelberg games are **not** limited to the single-leader single-follower case
- In a single-leader multi-follower case, the Stackelberg equilibrium is basically composed of an **optimal** policy for the leader with respect to a **Nash equilibrium** of the followers
  - It is often desirable to have a **unique** Nash equilibrium for the followers game, so as to make the Stackelberg solution tractable
  - **Example application**: Pricing for Internet Service Providers
- Multi-leader multi-follower Stackelberg games
  - At the Stackelberg equilibrium, both leaders and followers are in a Nash equilibrium (the Nash equilibria are correlated)
  - Hard to solve when the followers game has many equilibria

Example: Buyer/Seller Game (Two Level)

- **Buyer/Seller (Leader/Follower) Game**
  - Sender (buyer) buying the services from the relays to improve its performance, such as the transmission rate
  - Relays (sellers) selling service, such as power, by setting prices
  - Tradeoffs: Price too high, sender buying from others; price too low, profit low; sender decides to buy whose and how much to spend
  - Procedures: Convergence to the optimal equilibrium
  - Example: Power Control and Relay Section for Cooperative Transmission

$1000
Per Power

$800
Per Power

Example: Access Game

- The technique of **backward induction** helps to identify the Stackelberg equilibrium
  - It helps reducing the strategy space but becomes very complex for longer extensive form games
- **Definition:** The strategy profile **s** is a Stackelberg equilibrium with player *p1* as the *leader* and player *p2* as the follower
  - If player *p1* maximizes his payoff subject to the constraint that player *p2* chooses according to his best response function.
- Application:
  - If *p1* chooses *T* then the best response of *p2* is to play *Q* (payoff of 1-c)
  - If *p1* chooses *Q*, then the best response of *p2* is *T* (payoff of 0 for p1)
  - *p1* will therefore choose *T* which is the Stackelberg equilibria

p_1	p_2	payoff
T	T	(-c,-c)
T	Q	(1-c,0)
Q	T	(0,1-c)
Q	Q	(0,0)

Correlated Equilibrium (1)

- Beyond the Nash equilibrium, one can seek a more generalized solution concept for noncooperative games: **the correlated equilibrium**
- A concept suited for scenarios that involve a decision process in between non-cooperation and cooperation
- A correlated equilibrium is, in essence, a generalization of the Nash equilibrium
  - Requires an arbitrator who can send (private or public) signals to the players
  - These signals allow the players to coordinate their actions and perform joint randomization over strategies
- The arbitrator can be a virtual entity (the players can agree on the first word they hear on the radio) and generate signals that do not depend on the system

Correlated Equilibrium (2)

- A multi-strategy obtained using the signals is a set of strategies
  - One strategy for each player which may depend on all the information available to the player including the signal it receives
- It is said to be a correlated equilibrium if no player has an incentive to deviate unilaterally from its part of the multi-strategy
- A special type of "deviation" can be of course to ignore the signals!

Definition 22 *Given a strategic game G, a correlated strategy $p(\mathbf{s}) = p(s_i, \mathbf{s}_{-i})$ is said to be a* correlated equilibrium *if, for all $i \in \mathcal{N}$, $s_i, s_i' \in \mathcal{S}_i$, and $\mathbf{s}_{-i} \in \mathcal{S}_{-i}$, we have:*

$$\sum_{\mathbf{s}_{-i} \in \mathcal{S}_{-i}} p(s_i, \mathbf{s}_{-i}) \left[u_i(s_i', \mathbf{s}_{-i}) - u_i(s_i, \mathbf{s}_{-i}) \right] \leq 0. \tag{3.37}$$

Example: Correlated Equilibrium

- Each user to consider the joint distribution of users' actions

Table 1.1: Two secondary users game (from left to right): (a) payoff table; (b) Nash Equilibrium; (c) Mixed Nash Equilibrium; (d) Correlated Equilibrium.

	0	1
0	(5,5)	(6,3)
1	(3,6)	(0,0)

	0	1
0	0	(0 or 1)
1	(1 or 0)	0

	0	1
0	9/16	3/16
1	3/16	1/16

	0	1
0	0.6	0.2
1	0.2	0

Definition 11 (Correlated Equilibrium) *A probability distribution p is a correlated equilibrium of game* $G = [\mathcal{M}, \{\mathcal{A}_i\}_{i\in\mathcal{M}}, \{R_i\}_{i\in\mathcal{M}}]$, *if and only if, for all* $i \in \mathcal{M}$, $a_i \in \mathcal{A}_i$, *and* $a_{-i} \in \mathcal{A}_{-i}$,

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} p(a_i, a_{-i})[s_i(a_i', a_{-i}) - s_i(a_i, a_{-i})] \le 0, \forall a_i' \in \mathcal{A}_i.$$

By dividing inequality in (1.12) with $p(a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}} p(a_i, a_{-i})$, *we have*

$$\sum_{a_{-i}\in\mathcal{A}_{-i}} p(a_{-i}|a_i)[u_i(a_i', a_{-i}) - s_i(a_i, a_{-i})] \le 0, \forall a_i' \in \mathcal{A}_i.$$

- Multiple access game
  - Distributive Opportunistic Spectrum Access for Cognitive Radio using Correlated Equilibrium and No-regret Learning, WCNC 2007

Supermodular Games (1)

- Strategic complementarities: If a player chooses a higher action, the other players are better off if they also take a higher action
- Nice properties: Existence and achievability of NE

Definition 4 (Sublattice) *A real i-dimensional set $\mathcal{V}$ is a sublattice of $\Re^i$ if for any two elements $a, b \in \mathcal{V}$, the component-wise minimum (i.e., $a \wedge b$) and the component-wise maximum (i.e., $a \vee b$) are also in $\mathcal{V}$. In particular, a compact sublattice has a (component-wise) smallest and largest element. Any compact (one-dimensional) interval is a sublattice of $\mathbb{R}$.*

Definition 5 (Function with increasing differences) *A twice differentiable function f has increasing differences in variables (x, t) if $\partial^2 f/\partial x \partial t \geq 0$ for any feasible x and t.*[4]

Definition 6 (Supermodular function) *A function f is supermodular in $\boldsymbol{x} = (x_1, .., x_i)$ if it has increasing differences in (x_i, x_j) for all $i \neq j$.*

Definition 7 (Supermodular game) *A game $G = \left[\mathcal{M}, \{\mathcal{A}_i\}_{i \in \mathcal{M}}, \{s_i\}_{i \in \mathcal{M}}\right]$ is supermodular if for each player $i \in \rangle$, (a) the strategy space $\mathcal{P}_i$ is a nonempty and compact sublattice, and (b) the payoff function s_i is continuous in all players' strategies, is supermodular in player i's own strategy, and has increasing differences between any component of player i's strategy and any component of any other player's strategy.*

Supermodular Games (2)

Theorem 4 *In a supermodular game $G = \left[\mathcal{M}, \{\mathcal{A}_i\}_{i\in\mathcal{M}}, \{s_i\}_{i\in\mathcal{M}}\right]$,*

(a) *The set of NEs is a nonempty and compact sublattice and so there is a component-wise smallest and largest NE.*

(b) *If the users' best responses are single-valued, and each user uses MBR updates starting from the smallest (largest) element of its strategy space, then the strategies monotonically converge to the smallest (largest) NE.*

(c) *If each user starts from any feasible strategy and uses MBR updates, the strategies will eventually lie in the set bounded component-wise by the smallest and largest NE. If the NE is unique, the MBR updates globally converge to that NE from any initial strategies.*

- Power control in CDMA networks can often be captured using a supermodular game model
  - Dog barking effect

Wardrop Equilibrium (1)

- Wardrop (1952) postulated that users in a network game select routes of minimal length.
  - In the asymptotic regime, the game becomes a non-atomic one, in which the **impact of a single player on the others is negligible.**
- In the networking game context, the related solution concept is often called Wardrop equilibrium and is often much easier to compute than the original Nash equilibrium.
- The Wardrop equilibrium state can be seen as a state where no arbitrary small fraction of the traffic assigned to some path can benefit from unilaterally deviating to another path.
- Useful for games dealing with network flows and in which there exists a large population of players.
  - A. Haurie and P. Marcotte, "On the Relationship between Nash-Cournot and Wardrop Equilibria," Networks, vol. 15, pp. 295-308, 1985.

Wardrop Equilibrium (2)

Definition 28 *A feasible flow vector f is at* Wardrop equilibrium *if for every commodity $i \in [k]$ and paths $p_1, p_2 \in \mathcal{P}_i$ with $f_{p_1} > 0$, it holds that:*

$$l_{p_1}(f) \leq l_{p_1}(f). \tag{3.50}$$

- In the Wardrop equilibrium, a flow vector is considered stable when no fraction of the flow can improve its sustained latency by moving unilaterally to another path.
  - **Wardrop's first principle:** All used paths from a source to a destination have equal mean latencies.
  - **Wardrop's second principle:** Any unused path from a source to a destination has greater potential mean latency than that along the used paths.
- Mainly used in flow control and routing in communication networks, in dense ad hoc networks, as well as routing in wireless networks.

Summary

- Non-cooperative game is the most **basic** form of the game theory
- How to carefully design the utility function
  - Basic components
  - Convergence and uniqueness
- Extensive form
- Different types of games and their examples
- Price of Anarchy
  - Further reading: Pricing to improve the performance

    Cem Saraydar and Narayan B. Mandayam and David J. Goodman Efficient Power Control via Pricing in Wireless Data Networks IEEE Trans. on Communications, vol. 50, No. 2, pp. 291-303, February 2002