Lecture 24: October 31

Instructor: Ankur A. Kulkarni *Scribes: Sourabh, Pushpendra, Nachiket, Ishan, Chaitanya Pande*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

24.1 Game tree model vs State space approach

The extensive model of a game that we've discussed so far lays out the description in the layout of a tree. Actions taken by the players dictate which path actually gets executed. The system behaviour could effectively be summarized as an input-output model with player strategies as inputs and the payoffs at the end of the game considered as output. The information structure of the game also indicates whether the inputs(strategies) have a feedback component or not.

In the domain of control theory though, systems typically have a richer structure than an input-output mapping. The state-space abstraction assigns a *state* to the system at every time instant. The notion of a *state* encapsulates some sort of information about the configuration of the system. The evolution of the system then is a function of the current state and input(may or may not involve feedback) and so are the observations made about it. For example, the human body as a system can be described by variables like heart rate, body temperature and blood pressure. To know the state of the human body at some instant refers to determining the values of these variables.

This discussion induces a few questions like - whether there always is a state-space model corresponding to a game in strategies, and if true how does one convert one model of the system into the other or is the equivalence only valid under certain stringent assumptions.

The difference in the two approaches seemingly comes down to the notion of *causality*. In an extensive game, the order in which interactions occur between players is often fuzzy. To the contrary, in a state space description, past actions decide the current state which then determines the action taken at the current instant. There is a definite causal relation between the decisions made unlike a game played with strategies.

24.2 General formulation of Dynamic games

A dynamic game is specified with the following components:

1. A set of players, $\mathcal{N} = \{1, 2, \dots N\}$.
2. Set of instants at which an action is taken, $\mathcal{K} = \{1, 2, \dots k\}$.

3. An underlying state space, $\mathcal{X}$.
4. Sets of actions available to the players, $\mathcal{U}_k^i$.
5. State equation, $f_k : \mathcal{X} \times \mathcal{U}_k^1 \times \dots \mathcal{U}_k^N \rightarrow \mathcal{X}$.
6. Sets of observations made by players, Y_k^i.
7. Information sets of the players at different time instants, denoted as η_k^i.
   $\eta_k^i \in \{y_1^1, \dots y_k^1, \dots y_1^N \dots y_k^N, u_1^1 \dots, \dots u_k^N\}$
8. The payoffs of course, $\mathcal{L}^i : \mathcal{X} \times \mathcal{U}_k^1 \times \dots \mathcal{U}_k^N \rightarrow \mathbb{R}$
9. The actions taken by the players, a function of their information content at that stage.
   $u_k^i = \gamma_k^i(\eta_k^i) \in \mathcal{U}_k^i$, $\quad \Gamma_k^i =$ the class of all possible strategies(γ_k^i).

Throughout the above description, the superscripts denote the index of the player while the subscript refers to the time instant. Next we consider a few specific instances of dynamic games.

Case I: N = 1
The dynamic game would essentially be a Dynamic programming(Optimal control) problem for the lone player.

Case II: $\mathcal{L}^i = \mathcal{L}^j$ for all (i, j)
Problems where all players share the same payoff function are referred to as $Team$ *problems*. For instance, the operation of a power grid where the action-makers are sensory units, power generators or command stations and the common goal is to enhance some performance metric like reliability. Such games in general can't be modeled as a single agent optimization problem since the agents have different information sets. The effectiveness of the optimal solution depends to a large extent on the information structure of the game. Note however that a common payoff needn't make the game *cooperative*. That term is reserved for games where players can enter into pre-game contracts. Neither does a *team* problem imply that the agents have the same information.

24.2.1 Solution Concepts for state based games

Definition 24.2.1 (Person by person optimal solution). For the ith player, we define

$$J^i(\gamma^i, \gamma{-}i) = \mathcal{L}^i(u^1 = \gamma^1(\eta^1), \dots u^N = \gamma^N(\eta^N))$$

A person-by-person(*pbp*)-optimal solution is the set of strategies $(\gamma^{1*}, \dots \gamma^{N*})$ such that,

$$J^i(\gamma^{i*}, \gamma^{-i*}) \leq J^i(\gamma^i, \gamma^{-i*}) \forall \gamma^i$$

Definition 24.2.2 (Team optimal solution). Considering the set of games where all players have the same payoff function, if

$$J(\gamma^{1*}, \dots \gamma^{N*}) \leq J(\gamma^1, \dots \gamma^N) \quad \forall (\gamma^1, \dots \gamma^N)$$

then $(\gamma^{1*}, \dots \gamma^{N*})$ is a team optimal solution to the game.

It is evident that a team optimal solution in a common objective game is also a person-by-person optimal solution. The converse however doesn't always hold true. In fact, we have a rather specific result.

Theorem 24.1. *Consider a static simultaneous move team problem. If the payoff function J is differentiable and convex, then every person-by-person optimal solution is team optimal.*

We pick up a few examples illustrating the non-equivalence when the above conditions don't hold.

Example 24.2.1. Consider a 2-player single move game with the common payoff function specified as,

$$J(u^1, u^2) = (u^1)^2 + (u^2)^2 + 10u^1u^2 + 2u^1 + 3u^2$$

where the inputs u^1 and u^2 are the values chosen by the two players.

This game doesn't have a team optimal solution. For a given value of u^1, if the second player chooses $u^2 = \frac{-2}{3} \times u^1$, the payoff will be

$$J(u^1, u^2) = \frac{-47}{9} \times (u^1)^2$$

which approaches $-\infty$ for large values of u^1.

Whereas a *pbp* optimal solution can be found by taking partial derivatives of the payoff function wrt to u^1 and u^2, that gives,

$$2u^1 + 10u^2 + 2 = 0$$
$$2u^2 + 10u^1 + 3 = 0$$

The pair of linear equations gives $(\frac{-13}{48}, \frac{-7}{48})$ as a *pbp*-optimal solution.

Example 24.2.2. Consider another game where the common payoff function is non-differentiable.

$$J(u^1, u^2) = \begin{cases} (1 - u^2)^2 + (u^1)^2 & u^1 > u^2 \\ (1 - u^1)^2 + (u^2)^2 & u^1 \leq u^{2]} \end{cases}$$

The payoff function specified above is strictly convex and hence has a unique team optimal solution. Though non-differentiability renders it to have multiple *pbp*-optimal solutions.