15-888 Computational Game Solving (Fall 2021) Tue, Nov 16th 2021

Lecture 20

Correlated strategies and team coordination

Instructor: Gabriele Farina*

So far, we have focused on the task of computing optimal strategies for individual agents that seek to maximize their *own* utility. On the other hand, many realistic interactions require studying *correlated* strategies. As an example, consider a team of players that collude at a poker table: the optimal strategy for the team is to *strategically coordinate their moves*, and as a result the colluding players will not play according to independent strategies. The study of the geometric and analytical properties of correlated strategies in extensive-form strategic interactions is a fundamental question with far-reaching implications, and is yet to be fully explored.

1 Three models of team coordination

In this lecture, we will focus on the case of a team of two players coordinating against a third player. Depending on the amount of communication allowed among the team members, we can identify at least three solution concepts.

- If the team members can freely (and privately) communicate during play, the team effectively becomes a single player. Hence, all the tools we've seen so far (for example, learning an optimal strategy using CFR or linear programming) directly apply. We will refer to this equilibrium as 'Team Nash equilibrium'.
- If the team members cannot communicate at all, then the strategies of the team members should be picked as the pair of strategies that maximizes the expected utility of the team against a best-responding agent. This solution concept is called a *team maxmin equilibrium (TME)*. As we show in Section 2, the minmax theorem does not hold in general for TME. So, perhaps, the term "equilibrium" should be used carefully when referring to TME.
- The third case is intermediate, and models the setting where the team members have an opportunity to discuss and agree on tactics before the game starts, but are otherwise unable to communicate during the game, except through their publicly-observed actions. This models, for example, multi-player poker in the presence of collusion, and the game of bridge.

  More concretely, you can visualize TMECor as the following process: before playing, the team members get together in secret, and discuss about tactics for the game. They come up with m possible plans, each of which specifies a deterministic strategy for each team members, and write them down in m separate envelopes. Then, just before the game starts, they pick one of the m envelopes according to a shared probability distribution, and play according to the plan in the chosen envelope.

  It is important to realize that the sampling of the envelope can happen even if the team members cannot communicate before the game starts, as long as they can agree on some shared signal, such as for example a common clock. With that signal as input, the team members could seed a random number generator and use that to agree on a random envelope, without having communicated among each other.

*Computer Science Department, Carnegie Mellon University. ✉ gfarina@cs.cmu.edu.

Table 1 compares the properties of the three solution concepts defined above. Generally speaking, as the amount of allowed communication increases (from TME to TMECor to Team Nash), the utility of the team increases, while the complexity of computing the optimal strategy for the team decreases. In particular, as the first three rows of Table 1 show, the problem of computing a TME strategy for the team is significantly less well-behaved than TMECor.

	TME no communication *ever*	**TMECor** no communication *during play*	**Team Nash Eq.** private communication during play
Convex problem	✗	✓	✓
Bilinear saddle-point problem	✗	✓	✓
Low-dimensional strat. polytope	✗	✓	✓
Minmax theorem	✗	✓	✓
Team utility	low	higher	highest
Complexity	very hard	sometimes hard	polynomial

Table 1: Comparison between TME, TMECor and Team Nash equilibrium.

2 The minmax theorem fails for TME

As mentioned in the previous section, one of the major shortcomings of TME (at least from a game-theoretic point of view) is that team maxmin equilibria might fail the minmax theorem. We will now show an example of this.

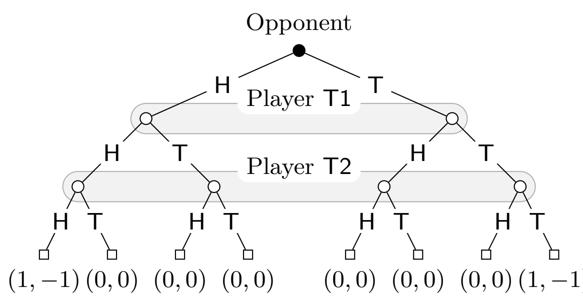

Figure 1: Matching Pennies game.

Consider the matching pennies game of Figure 1 (Left), where a team of two players (denoted T1 and T2 respectively) compete against an opponent. Each player has a penny. First, the opponent secretly turns their penny heads or tails. Then, Player T1 secretly turns their penny heads or tail. Finally, Player T2 turns their penny heads or tail. After all pennies have been turned, all of them are revealed. If all three pennies match (that is, they are all heads or they are all tails), the team wins a payoff of 1 and the opponent suffers a loss of −1. Otherwise, nobody gains or loses anything.

Maxmin value The TME strategy for the team is the solution to the *maxmin* problem

$$v_{\text{maxmin}} \coloneqq \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & \begin{cases} \min\limits_{\boldsymbol{z}} & x_{\mathsf{H}} \cdot y_{\mathsf{H}} \cdot z_{\mathsf{H}} + x_{\mathsf{T}} \cdot y_{\mathsf{T}} \cdot z_{\mathsf{T}} \\ \text{s.t.} & \textcircled{1}\ z_{\mathsf{H}} + z_{\mathsf{T}} = 1 \\ & \textcircled{2}\ z_{\mathsf{H}}, z_{\mathsf{T}} \ge 0 \end{cases} \\ \text{s.t.} & \textcircled{3}\ x_{\mathsf{H}} + x_{\mathsf{T}} = 1 \\ & \textcircled{4}\ y_{\mathsf{H}} + y_{\mathsf{T}} = 1 \\ & \textcircled{5}\ x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}} \ge 0. \end{cases} \qquad (\spadesuit)$$

Lemma 2.1. The solution to the maxmin problem defined in (♠) is $v_{\text{maxmin}} = 1/4$.

Proof. The internal minimization problem in (♠) is minimizing a linear function over the 2-simplex. So, the solution to the internal minimization problem is the minimum of the function over the two vertices, that is, $\min\{x_{\mathsf{H}} \cdot y_{\mathsf{H}}, y_{\mathsf{T}} \cdot y_{\mathsf{T}}\}$. Hence, we can rewrite (♠) as

$$v_{\text{maxmin}} = \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & \min\{x_{\mathsf{H}} \cdot y_{\mathsf{H}}, y_{\mathsf{T}} \cdot y_{\mathsf{T}}\} \\ \text{s.t.} & \textcircled{3}\ x_{\mathsf{H}} + x_{\mathsf{T}} = 1 \\ & \textcircled{4}\ y_{\mathsf{H}} + y_{\mathsf{T}} = 1 \\ & \textcircled{5}\ x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}} \ge 0. \end{cases}$$

We can now perform the substitutions $x_{\mathsf{T}} = 1 - x_{\mathsf{H}}$ and $y_{\mathsf{T}} = 1 - y_{\mathsf{H}}$ (justified by ③ and ④), and get to

$$v_{\text{maxmin}} = \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & \min\{x_{\mathsf{H}} \cdot y_{\mathsf{H}}, (1 - x_{\mathsf{H}}) \cdot (1 - y_{\mathsf{H}})\} \\ \text{s.t.} & \textcircled{5}\ 0 \le x_{\mathsf{H}} \le 1, \quad 0 \le y_{\mathsf{H}} \le 1. \end{cases}$$

Not that the problem is completely symmetric: if $(x^*_{\mathsf{H}}, y^*_{\mathsf{H}})$ is an optimal solution, then $(1 - x^*_{\mathsf{H}}, 1 - y^*_{\mathsf{H}})$ is also optimal. Hence, there exists at least one optimal solution in which

$$x_{\mathsf{H}} \cdot y_{\mathsf{H}} \le (1 - \mathsf{H}) \cdot (1 - y_{\mathsf{H}}) \qquad \iff \qquad x_{\mathsf{H}} + y_{\mathsf{H}} \le 1$$

and consequently

$$v_{\text{maxmin}} = \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & x_{\mathsf{H}} \cdot y_{\mathsf{H}} \\ \text{s.t.} & \textcircled{6}\ x_{\mathsf{H}} + y_{\mathsf{H}} \le 1 \\ & \textcircled{5}\ 0 \le x_{\mathsf{H}} \le 1, \quad 0 \le y_{\mathsf{H}} \le 1. \end{cases}$$

Since x_{H} is nonnegative, the objective is maximized when y_{H} is maximized, which happens for $y_{\mathsf{H}} = 1 - x_{\mathsf{H}}$. So,

$$v_{\text{maxmin}} = \begin{cases} \max\limits_{x_{\mathsf{H}}} & x_{\mathsf{H}} \cdot (1 - x_{\mathsf{H}}) \\ \text{s.t.} & \textcircled{5}\ 0 \le x_{\mathsf{H}} \le 1, \end{cases}$$

which is the maximum of a one-dimensional parabola. Taking the gradient of the objective, it's immediate to check that the maximum is obtained when $x_{\mathsf{H}} = 1/2$, and hence the statement of the lemma follows. □

Minmax value Conversely, the TME strategy for the opponent is the solution to the *maxmin* problem

$$v_{\text{minmax}} := \begin{cases} \min\limits_{\boldsymbol{z}} & \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & x_{\mathsf{H}} \cdot y_{\mathsf{H}} \cdot z_{\mathsf{H}} + x_{\mathsf{T}} \cdot y_{\mathsf{T}} \cdot z_{\mathsf{T}} \\ \text{s.t.} & \textcircled{1} \;\; x_{\mathsf{H}} + x_{\mathsf{T}} = 1 \\ & \textcircled{2} \;\; y_{\mathsf{H}} + y_{\mathsf{T}} = 1 \\ & \textcircled{3} \;\; x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}} \ge 0. \end{cases} \\ \text{s.t.} & \textcircled{4} \;\; z_{\mathsf{H}} + z_{\mathsf{T}} = 1 \\ & \textcircled{5} \;\; z_{\mathsf{H}}, z_{\mathsf{T}} \ge 0 \end{cases} \tag{$\clubsuit$}$$

Lemma 2.2. The solution to the minmax problem defined in ($\clubsuit$) is $v_{\text{minmax}} = 1/2$.

Proof. The choices $(x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}}) = (1, 0, 1, 0)$ and $(x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}}) = (0, 1, 0, 1)$ are feasible for the internal maximization problem in ($\clubsuit$). Hence, we have

$$\begin{aligned} v_{\text{minmax}} &\ge \begin{cases} \min\limits_{\boldsymbol{z}} & \max\{z_{\mathsf{H}}, z_{\mathsf{T}}\} \\ \text{s.t.} & \textcircled{4} \;\; z_{\mathsf{H}} + z_{\mathsf{T}} = 1 \\ & \textcircled{5} \;\; z_{\mathsf{H}}, z_{\mathsf{T}} \ge 0 \end{cases} \\ &= \begin{cases} \min\limits_{\boldsymbol{z}} & \max\{z_{\mathsf{H}}, 1 - z_{\mathsf{H}}\} \\ \text{s.t.} & \textcircled{5} \;\; 0 \le z_{\mathsf{H}} \le 1. \end{cases} \end{aligned}$$

It's immediate to see that the solution to the one-dimensional optimization problem on the right-hand side is $1/2$, attained for $z_{\mathsf{H}} = 1/2$. So $v_{\text{minmax}} \ge 1/2$. To complete the proof, it is enough to show that the there exists an assignment of $\boldsymbol{z}$ for which ($\clubsuit$) attains the value $1/2$.

To that end, consider the feasible assignment $z_{\mathsf{H}} = z_{\mathsf{T}} = 1/2$. Then, the objective of ($\clubsuit$) is

$$\begin{aligned} &\begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & 1/2 \cdot x_{\mathsf{H}} \cdot y_{\mathsf{H}} + 1/2 \cdot x_{\mathsf{T}} \cdot y_{\mathsf{T}} \\ \text{s.t.} & \textcircled{1} \;\; x_{\mathsf{H}} + x_{\mathsf{T}} = 1 \\ & \textcircled{2} \;\; y_{\mathsf{H}} + y_{\mathsf{T}} = 1 \\ & \textcircled{3} \;\; x_{\mathsf{H}}, x_{\mathsf{T}}, y_{\mathsf{H}}, y_{\mathsf{T}} \ge 0. \end{cases} \\ &= \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & 1/2(x_{\mathsf{H}} \cdot y_{\mathsf{H}} + (1 - x_{\mathsf{H}})(1 - y_{\mathsf{H}})) \\ \text{s.t.} & \textcircled{3} \;\; 0 \le x_{\mathsf{H}} \le 1, \quad 0 \le y_{\mathsf{H}} \le 1. \end{cases} \\ &= \begin{cases} \max\limits_{\boldsymbol{x},\boldsymbol{y}} & (x_{\mathsf{H}} - 1/2)(y_{\mathsf{H}} - 1/2) + 1/4 \\ \text{s.t.} & \textcircled{3} \;\; 0 \le x_{\mathsf{H}} \le 1, \quad 0 \le y_{\mathsf{H}} \le 1. \end{cases} \\ &= 1/2, \end{aligned}$$

where the last equality follows from since the product in the objective is maximized for $x_{\mathsf{H}} = y_{\mathsf{H}} = 1$. □

3 Modeling TMECor

As mentioned before, we can model a TMECor as a distribution μ over the product of deterministic strategies $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$ of the team members. Before the game starts, the team members will sample a pair of deterministic $(\pi_{\mathsf{T1}}, \pi_{\mathsf{T2}}) \in \Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$ and will simply play the deterministic strategy. Of course, μ will be chosen as the the distribution that guarantees the highest possible expected utility against a best-responding opponent, that is, as the solution to

$$\max_{\mu \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})} \min_{\boldsymbol{q}_{\mathsf{opp}} \in Q_{\mathsf{opp}}} \mathbb{E}_{(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}) \sim \mu}[u_{\mathsf{team}}(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}, \boldsymbol{q}_{\mathsf{opp}})]. \tag{1}$$

As usual, we can expand the definition of the utility of the game as a sum over all terminal nodes $z \in Z$ of the game, as

$$\begin{aligned}\mathbb{E}[u_{\mathsf{team}}(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}, \boldsymbol{q}_{\mathsf{opp}})] &= \sum_{(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}})} \mu[\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}] \left(\sum_{z \in Z} u_{\mathsf{team}}(z) \cdot \boldsymbol{\pi}_{\mathsf{T1}}[\sigma_{\mathsf{T1}}(z)] \cdot \boldsymbol{\pi}_{\mathsf{T2}}[\sigma_{\mathsf{T2}}(z)] \cdot \boldsymbol{q}_{\mathsf{opp}}[\sigma_{\mathsf{opp}}(z)] \cdot p_{\mathsf{chance}}(z) \right) \\ &= \sum_{z \in Z} u_{\mathsf{team}}(z) \cdot p_{\mathsf{chance}}(z) \cdot \left(\sum_{(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}})} \mu[\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}] \cdot \boldsymbol{\pi}_{\mathsf{T1}}[\sigma_{\mathsf{T1}}(z)] \cdot \boldsymbol{\pi}_{\mathsf{T2}}[\sigma_{\mathsf{T2}}(z)] \right) \cdot \boldsymbol{q}_{\mathsf{opp}}[\sigma_{\mathsf{opp}}(z)].\end{aligned}$$

Hence, we immediately have the following.

Observation 3.1. The above expression for the expected utility of the team is bilinear in μ and $\boldsymbol{q}_{\mathsf{opp}}$. This shows that the computation of a TMECor is a bilinear saddle-point problem (and therefore in particular it can be converted into a linear program, which is itself a convex problem, cf. Table 1). Another consequence of the bilinear saddle-point structure is the fact that the minmax theorem must hold for TMECor strategies, unlike TME (as we have seen in Section 2).

However, at this point it is still unclear how to rewrite the problem so that the set of strategies for the team is a polytope belonging to a low-dimensional space.

Realization Form Representation The expression for the expected utility of the team derived above depends on the distribution of play μ only as a linear function of the aggregate quantities

$$\boldsymbol{\omega}[z] := \sum_{(\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}})} \mu[\boldsymbol{\pi}_{\mathsf{T1}}, \boldsymbol{\pi}_{\mathsf{T2}}] \cdot \boldsymbol{\pi}_{\mathsf{T1}}[\sigma_{\mathsf{T1}}(z)] \cdot \boldsymbol{\pi}_{\mathsf{T2}}[\sigma_{\mathsf{T2}}(z)], \qquad z \in Z. \tag{2}$$

Hence, at least conceptually, we can operate a change of variables in the optimization problem, trading the distribution μ, for the vector $\boldsymbol{\omega}$. Doing so greatly reduces the number of variables in the optimization problem: μ requires $\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}}$ scalar variables (an exponential amount in the game tree size) to be specified, while $\boldsymbol{\omega}$ only requires $|Z|$ scalar variables (a linear amount in the game tree size).

In order to perform the change of variable effectively, we need to study the new domain of the optimization problem. In particular, as we move away from $\mu \in \Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})$ and allow the team to specify a $\boldsymbol{\omega}$ instead, we need to characterize the domain of $\boldsymbol{\omega}$ that can be induced from μ according to (2). The key observation here is that the mapping from μ to $\boldsymbol{\omega}$ is *linear*. Given that the set of feasible μ is a convex polytope (the simplex $\Delta(\Pi_{\mathsf{T1}} \times \Pi_{\mathsf{T2}})$), we can use the following lemma to conclude that the set of all feasible $\boldsymbol{\omega}$ is itself a convex polytope.

Lemma 3.1. Let S be a convex polytope, and f be a linear transformation. Then, the image of S under f, that is, the set $f(S) := \{f(\boldsymbol{s}) : \boldsymbol{s} \in S\}$, is a convex polytope.

Proof. We will show that $f(S)$ is a convex polytope by showing that it is the convex combination of a finite set of points. Let $\{v_1, \ldots, v_m\}$ be the vertices of S—that is, $S = \text{co}\{v_1, \ldots, v_m\}$. We will show that $f(S) = \text{co}\{f(v_1), \ldots, f(v_m)\}$, by proving the two directions of the inclusion separately.

($\subseteq$) We start by arguing that $f(S) \subseteq \text{co}\{f(v_i), \ldots, f(v_m)\}$. Pick any $\boldsymbol{x} \in f(S)$; we will show that $\boldsymbol{x} \in \text{co}\{f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_m)\}$. By definition of $f(S)$, there must exist (at least one) $\boldsymbol{s} \in S$ such that $\boldsymbol{x} = f(\boldsymbol{s})$. Now, since S is a convex polytope with vertices $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, there exist convex combination coefficients $(\lambda_1, \ldots, \lambda_m) \in \Delta^m$ such that $\boldsymbol{s} = \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_m \boldsymbol{v}_m$. Using the linearity of f we can then write

$$\boldsymbol{x} = f(\boldsymbol{s}) = \lambda_1 f(\boldsymbol{v}_1) + \cdots + \lambda_m f(\boldsymbol{v}_m).$$

As the right-hand side is a convex combination of the points $f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_m)$, we conclude that $\boldsymbol{x} \in \operatorname{co}\{f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_m)\}$ as we wanted to show.

($\supseteq$) Next, we argue that $f(S) \supseteq \operatorname{co}\{f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_m)\}$. Pick any point $\boldsymbol{x} \in \operatorname{co}\{f(v_1), \ldots, f(\boldsymbol{v}_m)\}$; we will argue that $\boldsymbol{x} \in f(S)$. By definition of convex hull, there exist convex combination coefficients $(\lambda_1, \ldots, \lambda_m) \in \Delta^m$ such that $\boldsymbol{x} = \lambda_1 f(\boldsymbol{v}_1) + \cdots + \lambda_m f(\boldsymbol{v}_m)$. By linearity of f, we can therefore write $\boldsymbol{x} = f(\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_m \boldsymbol{v}_m)$. Since the $\boldsymbol{v}_i$'s are the vertices of the convex polytope S, the point $\boldsymbol{s} \coloneqq \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_m \boldsymbol{v}_m$ is a point in S, and therefore $\boldsymbol{x} = f(\boldsymbol{s}) \in f(S)$ as we wanted to show. $\square$