CSCE 631-D: Intelligent Agents: Computational Game Solving

Lecture 3: Nash Equilibria

Alan Kuhnle

Fall 2026

Texas A&M University

Today's Plan

Learning objectives. By the end of this lecture you should be able to:

1. Compute Nash equilibria in 2×2 games using the indifference principle.
2. State and interpret the Nash existence theorem.
3. Distinguish maxmin strategies from Nash equilibrium strategies.

Best Responses

Best Response

If you know what your opponents are doing, what should you play? The *best response* formalizes the answer: it is the strategy that maximizes your payoff given the opponents' choices.

Definition

Given a mixed-strategy profile σ_{-i} (the strategies of all players except i), a strategy σ_i^* is a **best response** to σ_{-i} if

$$u_i(\sigma_i^*, \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i}) \quad \forall\, \sigma_i \in \Delta(S_i).$$

(Here $\Delta(S_i)$ is the set of probability distributions over S_i, i.e., mixed strategies for player i.)

The **best-response set** (or correspondence) collects all optimal strategies:

$$BR_i(\sigma_{-i}) = \arg\max_{\sigma_i \in \Delta(S_i)} u_i(\sigma_i, \sigma_{-i}).$$

Key Observation

Since u_i is linear in σ_i, there always exists a **pure-strategy** best response. That is, $BR_i(\sigma_{-i}) \cap S_i \neq \emptyset$.

Best Response: Example

Consider Battle of the Sexes:

	Opera	Football
Opera	$(3, 2)$	$(0, 0)$
Football	$(0, 0)$	$(2, 3)$

Let Player 2 play Opera with probability q. Player 1's expected payoffs:

$$u_1(\text{Opera}, q) = 3q$$
$$u_1(\text{Football}, q) = 2(1 - q)$$

Best response of Player 1:

$$BR_1(q) = \begin{cases} \text{Opera} & \text{if } q > 2/5, \\ \text{Football} & \text{if } q < 2/5, \\ \Delta(\{O, F\}) & \text{if } q = 2/5. \end{cases}$$

Best Response Among Prompt Policies

[Agent-Thread Stop 1: Empirical Best Response Among Prompt Policies]

In the LLM agent setting, we apply the best-response concept by searching over a finite set of prompt policies. Fix the opponent's prompt policy, then find the one in your set with the highest estimated payoff.

Concrete example. Three candidate negotiation prompts vs. a fixed "cooperative" opponent (payoffs estimated from $N = 30$ sampled interactions each):

Your prompt policy	Est. payoff ($\hat{u}_1$)	95% CI
Aggressive	2.4	± 0.3
Reciprocal	3.1	± 0.2
Conciliatory	2.8	± 0.4

Empirical best response: **Reciprocal** ($\arg\max$ row).

Takeaway. This is the agent-world meaning of "best response" used all semester: a finite policy search over estimated payoffs.

Nash Equilibrium

Nash Equilibrium: Definition

A best response answers: "what should I do, given my opponents' strategies?" A **Nash equilibrium** answers the harder question: what happens when *every* player reasons this way simultaneously?

Definition (Nash, 1950)

A mixed-strategy profile $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ is a **Nash equilibrium** (NE) if for every player $i \in N$:

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*) \quad \forall\, \sigma_i \in \Delta(S_i).$$

Equivalently: $\sigma_i^* \in BR_i(\sigma_{-i}^*)$ for all i. In words, a NE is a **fixed point** of the best-response correspondence.

Key intuition: no player has an incentive to *unilaterally* deviate. Once reached, no single player can do better by changing only their own strategy.

Pure-Strategy Nash Equilibrium

The simplest NE to find are those where no player randomizes.

Definition

A **pure-strategy Nash equilibrium** (PSNE) is a strategy profile $s^* = (s_1^*, \ldots, s_n^*) \in S$ such that for all i and all $s_i \in S_i$:

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*).$$

Finding PSNE in a payoff matrix. For each cell, check whether each player's strategy is a best response to the other's.

Method: underline best-response payoffs. A cell where *all* players' payoffs are underlined is a PSNE.

Finding PSNE: Prisoner's Dilemma

	Cooperate	Defect
Cooperate	$(-1, -1)$	$(-3, \underline{0})$
Defect	$(\underline{0}, -3)$	$(\underline{-2}, \underline{-2})$

Underlined payoffs mark best responses.

- Row player's best response to Cooperate: Defect $(0 > -1)$.
- Row player's best response to Defect: Defect $(-2 > -3)$.
- Column player: symmetric.

Unique PSNE: (Defect, Defect) with payoff $(-2, -2)$.

Finding PSNE: Battle of the Sexes

	Opera	Football
Opera	$(\underline{3}, \underline{2})$	$(0, 0)$
Football	$(0, 0)$	$(\underline{2}, \underline{3})$

Two PSNE: (Opera, Opera) and (Football, Football).

Neither off-diagonal cell is a NE: each has a player who wants to switch.

Question: Is there also a *mixed* NE? Yes; we find it shortly.

Finding PSNE: Matching Pennies

	Heads	Tails
Heads	$(\underline{1}, -1)$	$(-1, \underline{1})$
Tails	$(-1, \underline{1})$	$(\underline{1}, -1)$

No cell has both payoffs underlined. **No PSNE exists.**

This is why we need mixed strategies: Nash's theorem guarantees a mixed NE exists.

Nash's Existence Theorem

Nash's Existence Theorem (1950)

Does every finite game have a Nash equilibrium? Not necessarily in pure strategies (Matching Pennies has none), but Nash proved that a *mixed-strategy* NE always exists.

Theorem (Nash, 1950)

Every finite game (finite players, finite strategies) has at least one **mixed-strategy Nash equilibrium**.

Proof idea (details on the next slide). Define the best-response correspondence $BR : \prod_i \Delta(S_i) \rightrightarrows \prod_i \Delta(S_i)$. The domain is compact and convex (product of simplices). BR is upper-hemicontinuous with convex values. By **Kakutani's fixed-point theorem**, BR has a fixed point, which is exactly a Nash equilibrium.

Note: This is a pure existence result. It says nothing about:

- How many NE exist.
- How to *compute* one efficiently.
- Which NE is "the right one" (equilibrium selection).

Kakutani's Fixed-Point Theorem

Kakutani's theorem is the mathematical engine behind Nash's existence proof. It generalizes the Brouwer fixed-point theorem from functions to set-valued maps (correspondences).

Theorem (Kakutani, 1941)

Let $X \subseteq \mathbb{R}^n$ be compact and convex, and let $F : X \rightrightarrows X$ be an upper-hemicontinuous correspondence with non-empty, convex values. Then F has a fixed point: $\exists\, x^* \in X$ with $x^* \in F(x^*)$.

Informally, upper-hemicontinuity means that small changes in the input produce only small changes in the output set. Failure of upper-hemicontinuity would mean the set of best responses suddenly shrinks as we nudge the opponent's strategy; limit points of best responses would be excluded.

Computing NE in 2×2 Games

The Indifference Principle

Nash's theorem guarantees that a mixed NE exists, but how do we *find* it? The indifference principle converts the existence question into a system of equations we can solve.

Theorem (Indifference Principle)

Let σ^* be a Nash equilibrium. For each player i, every pure strategy in $\mathrm{supp}(\sigma_i^*)$ (the set of pure strategies played with positive probability) yields the same expected payoff against σ_{-i}^*, and every pure strategy *outside* the support yields a weakly lower payoff.

Proof: Suppose $s_i, s_i' \in \mathrm{supp}(\sigma_i^*)$ with $u_i(s_i, \sigma_{-i}^*) > u_i(s_i', \sigma_{-i}^*)$. Then player i could increase payoff by shifting probability from s_i' to s_i, contradicting σ_i^* being a best response. $\square$

Takeaway: to compute a mixed NE, set up equations requiring each player to be

Support Enumeration Method

We will use this method on the 3×3 example at the end of this lecture and develop it formally in Module 2.

To find all NE of a two-player game:

Algorithm: Support Enumeration

1. For each pair of supports $(T_1 \subseteq S_1,\ T_2 \subseteq S_2)$:
   (a) Solve the indifference equations: for each player i, all pure strategies in T_i must yield equal payoff.
   (b) Check non-negativity: $\sigma_i(s_i) \geq 0$ for all $s_i \in T_i$.
   (c) Check best response: no strategy outside T_i yields strictly higher payoff.
2. Collect all valid solutions.

For a 2×2 game: at most $3 \times 3 = 9$ support pairs, but in practice just check the full-support case $(T_1 = S_1, T_2 = S_2)$ plus the PSNE. The full algorithm and its complexity are covered in M2 Lecture 5.

Example: Mixed NE in Battle of the Sexes

	Opera (q)	Football ($1-q$)
Opera (p)	$(3,2)$	$(0,0)$
Football ($1-p$)	$(0,0)$	$(2,3)$

Player 1 must be indifferent (given q):

$$3q = 2(1-q) \quad \Longrightarrow \quad q = \frac{2}{5}.$$

Player 2 must be indifferent (given p):

$$2p = 3(1-p) \quad \Longrightarrow \quad p = \frac{3}{5}.$$

Mixed NE: $\sigma_1 = (3/5, 2/5)$, $\sigma_2 = (2/5, 3/5)$.

Expected payoff: $u_1 = 3 \cdot 2/5 = 6/5$, $u_2 = 2 \cdot 3/5 = 6/5$.

Total: 3 Nash equilibria: two pure and one mixed.

Question: Both pure NE give one player their preferred outcome. Does the mixed NE split the surplus more fairly? What does each player's expected payoff of 6/5 tell you?

Example: Computing the Mixed NE of Matching Pennies

	H (q)	T ($1-q$)
H (p)	$(1,-1)$	$(-1,1)$
T ($1-p$)	$(-1,1)$	$(1,-1)$

Player 1 indifferent: $q-(1-q) = -(q)+(1-q) \Rightarrow 2q-1 = 1-2q \Rightarrow q = 1/2$.

Player 2 indifferent: $-p+(1-p) = p-(1-p) \Rightarrow 1-2p = 2p-1 \Rightarrow p = 1/2$.

Unique NE: $\sigma_1 = \sigma_2 = (1/2, 1/2)$, value $= 0$.

Notice: in the mixed NE, each player is *indifferent* between their pure strategies. The mixing probabilities are determined by the *opponent's* indifference condition.

How Many Nash Equilibria?

Theorem (Wilson, 1971; Harsanyi, 1973)
For "generic" games (payoffs in general position), the number of Nash equilibria is finite and **odd**.

Examples:

- Prisoner's Dilemma: 1 NE (pure).
- Matching Pennies: 1 NE (mixed).
- Battle of the Sexes: 3 NE (2 pure + 1 mixed).
- Coordination game: 3 NE.

Non-generic games can have continua of NE (e.g., games with payoff ties).

For $n \times n$ games, the maximum number of NE can be exponential in n.

Equilibria of Estimated Games

NE of an Estimated Game

[Agent-Thread Stop 2: Equilibrium Sensitivity to Payoff Estimation Error]

When payoffs come from sampled LLM interactions, equilibrium analysis rests on estimates, not exact values. The following table shows representative payoffs estimated from $N = 30$ interactions per cell:

	Cooperate	Defect
Cooperate	$(2.83 \pm 0.24,\ 2.83 \pm 0.24)$	$(0.47 \pm 0.30,\ 4.60 \pm 0.31)$
Defect	$(4.60 \pm 0.31,\ 0.47 \pm 0.30)$	$(1.07 \pm 0.12,\ 1.07 \pm 0.12)$

Point estimates: Defect is strictly dominant for both players. Unique NE: (Defect, Defect).

Sensitivity check. Give Cooperate its upper CI bound and Defect its lower CI bound (the most favorable scenario for Cooperate). Against C: upper-CI Cooperate 3.07 vs. lower-CI Defect 4.29; against D: 0.77 vs. 0.95. In both cases Defect still dominates, so (D, D) remains the unique NE throughout these confidence intervals.

PA1 core question: how sensitive are your equilibrium conclusions to payoff estimation error?

Maxmin Strategies

Maxmin Strategy

Nash equilibrium assumes all players are rational and play best responses. What if you cannot rely on that assumption?

The *maxmin strategy* answers this question: it optimizes against the **worst case**, providing a guaranteed payoff floor regardless of how the opponent behaves.

Definition

The **maxmin value** of player i is:

$$\overline{v}_i = \max_{\sigma_i \in \Delta(S_i)} \min_{\sigma_{-i} \in \Delta(S_{-i})} u_i(\sigma_i, \sigma_{-i}).$$

A strategy σ_i^* achieving this maximum is a **maxmin strategy** for player i.

Interpretation: The maxmin strategy maximizes player i's payoff under the worst-case opponent behavior.

Minmax Strategy

The maxmin value asks: what can player i guarantee? The *minmax* value asks the complementary question: what can the opponents force?

Definition
The **minmax value** against player i is:

$$\underline{v}_i = \min_{\sigma_{-i} \in \Delta(S_{-i})} \max_{\sigma_i \in \Delta(S_i)} u_i(\sigma_i, \sigma_{-i}).$$

$\underline{v}_i$ is the lowest payoff that the other players can *force* upon player i.

Proposition
For any game, $\overline{v}_i \leq \underline{v}_i$.

Proof: Choosing σ_i without knowing σ_{-i} cannot yield a higher guarantee than choosing after observing σ_{-i}, so max-min $\leq$ min-max. □

Notation mnemonic. The overbar $\overline{v}_i$ is the player's own guarantee: the payoff they can secure regardless of the opponent. It is always weakly lower than the underbar $\underline{v}_i$, which is what the opponent can hold the player to.

Equality holds for two-player zero-sum games (von Neumann).

Maxmin Example: Battle of the Sexes

Returning to Battle of the Sexes, which we have now analyzed through three different lenses: pure-strategy NE, mixed NE, and now the maxmin approach.

	Opera	Football
Opera	3	0
Football	0	2

Player 1 plays Opera with probability p. Worst-case payoff:

$$\min\{3p,\ 2(1-p)\} = \begin{cases} 3p & \text{if } p \leq 2/5, \\ 2(1-p) & \text{if } p \geq 2/5. \end{cases}$$

Maximize: the two expressions meet at $3p = 2(1-p) \Rightarrow p = 2/5$.

Maxmin value: $\overline{v}_1 = 3 \cdot 2/5 = 6/5$.

Maxmin strategy: $\sigma_1 = (2/5, 3/5)$.

Note: the maxmin strategy is *not* the same as the NE strategy $\sigma_1 = (3/5, 2/5)$ in general-sum games.

Minimax Theorem for Zero-Sum Games

Von Neumann's Minimax Theorem

The gap between maxmin and minmax values raises a natural question: when are they equal? For zero-sum games, the answer is always.

We showed that $\overline{v}_i \leq \underline{v}_i$ in any game. In zero-sum games, a remarkable simplification occurs: the two values are *equal*.

Theorem (von Neumann, 1928)

In any finite two-player zero-sum game with payoff matrix $A \in \mathbb{R}^{m \times n}$:

$$\max_{\sigma_1 \in \Delta_m} \min_{\sigma_2 \in \Delta_n} \sigma_1^\top A \sigma_2 \;=\; \min_{\sigma_2 \in \Delta_n} \max_{\sigma_1 \in \Delta_m} \sigma_1^\top A \sigma_2 \;=\; v^*.$$

The common value v^* is the **value of the game**.

Proof Sketch (via LP Duality)

This sketch uses LP duality. If you have not seen this, the key fact is that every linear program has a dual whose optimal value equals the primal optimum (strong duality).

Player 1 solves:

$$\max_{\sigma_1 \in \Delta_m} \min_{\sigma_2 \in \Delta_n} \sigma_1^\top A \sigma_2 = \max_{\sigma_1 \in \Delta_m} \min_{j \in [n]} (A^\top \sigma_1)_j$$

(the min over mixed strategies is achieved at a pure strategy).

This is equivalent to the LP:

$$\max \; v \quad \text{s.t.} \quad A^\top \sigma_1 \geq v \cdot \mathbf{1}, \; \sigma_1 \in \Delta_m.$$

Player 2's problem is the **dual LP**:

$$\min \; w \quad \text{s.t.} \quad A \sigma_2 \leq w \cdot \mathbf{1}, \; \sigma_2 \in \Delta_n.$$

By **strong LP duality**: $v^* = w^*$.

Consequences for Zero-Sum Games

Theorem

In a two-player zero-sum game, the following are equivalent:

1. (σ_1^*, σ_2^*) is a Nash equilibrium.
2. σ_1^* is a maxmin strategy and σ_2^* is a minmax strategy.
3. (σ_1^*, σ_2^*) is a saddle point: $\sigma_1^\top A \sigma_2^* \leq {\sigma_1^*}^\top A \sigma_2^* \leq {\sigma_1^*}^\top A \sigma_2$ for all σ_1, σ_2.

Key implications:

- All NE yield the *same payoff* (v^* for Player 1, $-v^*$ for Player 2).
- NE strategies are *interchangeable*: if (σ_1, σ_2) and (σ_1', σ_2') are both NE, so is (σ_1, σ_2').
- Computing NE = solving an LP = polynomial time.

NE vs. Maxmin in General-Sum Games

The Battle of the Sexes example showed that the maxmin strategy $\sigma_1 = (2/5, 3/5)$ differs from the NE strategy $\sigma_1 = (3/5, 2/5)$. This is not a coincidence: in general-sum games, NE and maxmin strategies differ in fundamental ways.

Property	Zero-sum	General-sum
NE payoff unique?	Yes	No
NE $=$ maxmin?	Yes	Not in general
NE interchangeable?	Yes	No
Computable in poly-time?	Yes (LP)	Unknown (PPAD-complete)

The minimax theorem makes zero-sum games the “easy” special case. Most of the computational difficulty in algorithmic game theory comes from general-sum games.

Properties of Nash Equilibria

NE and Pareto Optimality

We have seen that zero-sum games are computationally easy. We now ask a different question about NE quality: are they efficient?

A Nash equilibrium is stable, but is it *good*? Stability does not imply efficiency: Nash equilibria need not be Pareto-optimal.

Prisoner's Dilemma: (Defect, Defect) is the unique NE with payoff $(-2, -2)$, but (Cooperate, Cooperate) gives $(-1, -1)$, which Pareto-dominates it.

The **Price of Anarchy** quantifies this efficiency gap. The *social cost* of a strategy profile is the sum of all players' costs.

Price of Anarchy (Koutsoupias & Papadimitriou, 1999)

$$\text{PoA} = \frac{\text{worst NE social cost}}{\text{optimal social cost}}.$$

Multiple Equilibria and Selection

When multiple NE exist, which one is "correct"?

Equilibrium refinements:

- **Pareto-dominant NE:** select the NE that Pareto-dominates others (if one exists).
- **Risk-dominant NE:** select the "safer" NE (Harsanyi & Selten, 1988).
- **Focal points:** Schelling (1960), cultural or contextual salience.
- **Correlated equilibrium:** a convex relaxation that encompasses all NE (Lecture 4).
- **Trembling-hand perfect NE:** robust to small perturbations.

Which equilibrium do LLM agents select? Shared training data and similar architectures create implicit focal points (Schelling): LLM agents may coordinate on the same NE without explicit communication. An empirical selection question you can test in PA1's coordination game.

Existence vs. Computation

Nash's theorem is *topological*: Kakutani's fixed-point theorem guarantees existence. But existence says nothing about efficient computation.

Computational Complexity of NE

- **Two-player zero-sum:** solvable in poly-time (LP).
- **Two-player general-sum:** computing *any* NE is **PPAD-complete** (Chen & Deng, 2006; Daskalakis, Goldberg & Papadimitriou, 2006).
- *n*-**player:** also PPAD-complete (or harder).

PPAD: "Polynomial Parity Argument on Directed graphs." Captures problems where a solution is guaranteed to exist (by a parity argument) but may be hard to find.

Module 2 covers algorithms (support enumeration, Lemke–Howson, LP for zero-sum) and the precise complexity landscape.

Comprehensive Example: A 3×3 Game

A 3×3 Game

We now apply the full toolkit (best responses, PSNE search, indifference principle) to a larger game.

	L	C	R
T	$(3,1)$	$(0,2)$	$(1,0)$
M	$(1,2)$	$(2,1)$	$(3,0)$
B	$(0,3)$	$(1,0)$	$(2,2)$

Pause: examine the matrix and find all pure-strategy Nash equilibria before advancing.

Step 1: find PSNE. Underline best-response payoffs.

Player 1's BR: to L: T (3); to C: M (2); to R: M (3).

Finding Mixed NE in the 3×3 Game

Step 2: apply the indifference principle. To choose a support pair, first eliminate dominated strategies: B is strictly dominated by M for Player 1 ($1 > 0$, $2 > 1$, $3 > 2$), and R is strictly dominated by L for Player 2 ($1 > 0$, $2 > 0$, $3 > 2$). The remaining strategies form the candidate support $\{T, M\} \times \{L, C\}$.

Try support $\{T, M\}$ for Player 1 and $\{L, C\}$ for Player 2.

Player 1's indifference (support $\{T, M\}$, given Player 2 plays L with prob q, C with prob $1 - q$):

$$3q + 0(1-q) = 1 \cdot q + 2(1-q) \quad \Longrightarrow \quad 3q = 2 - q \quad \Longrightarrow \quad q = 1/2.$$

Player 2's indifference (support $\{L, C\}$, given Player 1 plays T with prob p):

$$1 \cdot p + 2(1-p) = 2p + 1(1-p) \quad \Longrightarrow \quad 2 - p = p + 1 \quad \Longrightarrow \quad p = 1/2.$$

Summary

- **Best response:** the optimal strategy given opponents' play; for LLM agents, a finite search over prompt policies with estimated payoffs.
- **Nash equilibrium:** mutual best response. Every finite game has at least one mixed NE (Nash, 1950).
- **Indifference principle:** in a NE, all strategies in the support yield equal payoff. Key tool for computing NE.
- **Maxmin and minimax:** security levels; equality in zero-sum games (von Neumann); diverge in general-sum games.
- **Computing NE:** poly-time for zero-sum (LP); PPAD-complete for general-sum. Algorithms in M2.
- **Estimated games:** equilibrium conclusions can be sensitive to payoff estimation error. Confidence intervals from sampled LLM play propagate into equilibrium structure.