Game Theory (ECO 311)

Lecture Notes

Feda Mohammadi

Professor: Dr. Jarrod Burgh

Berea College

Spring 2026

These notes are for personal academic study. Examples and definitions follow standard game theory notation.

Contents

Chapter 1: Introduction to Strategic Interaction

1.1 What is Game Theory?

Big idea

Game theory studies **interactive decisions**: situations where multiple players choose actions, and the outcome for each player depends on what *everyone* does.

Game theory builds models of situations where several entities (players) make choices, and the combination of their choices produces an outcome that affects all of them. Because different players can value the same outcome differently, game theory is also the study of **rational behavior** in these interactive settings. A central feature is **strategic interdependence**: my best action depends on what I think others will do.

Strategic interdependence

In many problems, you cannot choose well without thinking about others. This is what makes game theory different from standard single-person decision problems. An interaction has strategic interdependence **if each player's outcome depends on the actions of other players.**

Example 1: Competing coffee shops

Two coffee shops choose prices. If one shop lowers its price, it can attract customers away from the other shop. Each shop's profit depends on *both* prices, not just its own price. This is a game because each shop's best decision depends on what the other shop does.

Conflict and cooperation

Many strategic situations lie somewhere between pure conflict and pure cooperation. Some games look like competition (each player wants the other to do poorly), and some games allow mutually beneficial outcomes if players coordinate or cooperate.

Example 2: Roommates choosing quiet hours

Two roommates each prefer quiet time to study, but they also want flexibility. If both agree on quiet hours, both benefit. But if one ignores the agreement, they may gain personally while harming the other. This is a strategic situation mixing cooperation (agreement helps) and conflict (temptation to deviate).

A very short history (why the field exists)

Game theory has roots in early probability problems (for example, analyzing the probability of winning card games as play unfolds). Later work studied finite games with perfect information and strategic reasoning. Key foundational contributions include work by von Neumann (including the minimax theorem) and the book *Theory of Games and Economic Behavior* by von Neumann and Morgenstern (1944), which gave early formal descriptions of major classes of games and core principles of the theory. Nash later introduced the concept of strategic equilibrium and proved existence results that became central to modern game theory.

Where game theory is used

Game theory is heavily used in economics, but it is also widely applied in biology, computer science, political science, operations research, and other areas of mathematics.

Course focus

Games are often classified as **strategic (non-cooperative)** versus **coalitional (cooperative)**. But here, we mainly study **strategic (non-cooperative)** game theory.

How we represent a game: extensive form vs normal form

There are two common ways to describe the rules and structure of strategic interaction.

Extensive form (game as rules)

The extensive form describes the **rules of the game** explicitly: who moves when, what actions are available at each point, what information players have during play, and how play leads to an outcome.

A key concept in extensive-form games is a **strategy**. A strategy is a complete plan that specifies what a player would do at every situation where they might have to act.

Example 3: Strategy as a complete plan

In chess, a strategy is not just “move my queen now.” It is a full plan describing what you would do for every possible configuration you might face later. That is why a strategy is a complete mapping from situations to actions.

Normal form (game as a mathematical object)

In normal form, we list each player’s strategy set and specify a mapping from a profile of strategies to outcomes and payoffs. In this representation, an n-player game can be described by a map from a product of n spaces into $\mathbb{R}^n$ (one payoff for each player).

The extensive form is often simpler to work with and can lead to computations with lower complexity. The normal form is a universal and compact mathematical formulation once strategies and payoffs are specified.

1.2 Mathematical Review

Game theory relies heavily on mathematical structure. Before studying strategic interaction, we need to review the mathematical tools that will appear repeatedly throughout this course.

Functions

A function assigns a number to each element in a set. In economics and game theory, functions usually represent preferences, payoffs, or beliefs.

If

$$f : X \longrightarrow \mathbb{R},$$

then for each $x \in X$, the function assigns a real number $f(x)$.

Example: Payoff Function

Suppose a firm chooses quantity q. Its profit is

$$\pi(q) = 10q - q^2.$$

Here, q is the action and $\pi(q)$ is the payoff.
The function tells us how profitable each possible action is.

Maximization

A central idea in game theory is that players maximize something, usually utility or payoff.

Let $f(x)$ be a payoff function. A value x^* solves

$$x^* \in \arg\max_{x \in X} f(x)$$

if

$$f(x^*) \geq f(x) \quad \text{for all } x \in X.$$

This means x^* gives the highest attainable payoff.

Example: Simple Maximization

Let

$$f(x) = -x^2 + 4x.$$

Take the derivative:

$$f'(x) = -2x + 4.$$

Set it equal to zero:

$$-2x + 4 = 0 \quad \Rightarrow \quad x = 2.$$

So, $x = 2$ maximizes the function.

Sets and Strategies

Game theory describes choices using sets. If player i has a strategy set S_i, then:

$$S_i = \{\text{all strategies available to player } i\}.$$

If there are n players, the set of strategy profiles is

$$S = S_1 \times S_2 \times \cdots \times S_n.$$

An element of S is written

$$s = (s_1, s_2, \ldots, s_n).$$

Example: Two-Player Strategy Set

Suppose two firms choose either High price (H) or Low price (L).
Then

$$S_1 = \{H, L\}, \quad S_2 = \{H, L\}.$$

The strategy profiles are:

$$(H, H), (H, L), (L, H), (L, L).$$

Each profile represents one possible outcome.

Basic Probability

Many games involve uncertainty. A probability distribution assigns likelihoods to possible events.

If $E_1, E_2, \ldots, E_k$ are mutually exclusive events, then

$$\sum_{j=1}^{k} P(E_j) = 1.$$

Expected value plays a crucial role in decision making under uncertainty. If a random variable X takes values x_j with probabilities p_j, then

$$E[X] = \sum_j p_j x_j.$$

Example: Expected Payoff

Suppose a player receives

$$100 \text{ with probability } 0.5 \quad \text{and} \quad 0 \text{ with probability } 0.5.$$

Then the expected payoff is

$$E[X] = 0.5(100) + 0.5(0) = 50.$$

A rational decision maker compares expected payoffs across actions.

1.3 Some Basic Mathematics That Will Come in Handy Later

In solving maximization problems and analyzing curvature properties of payoff functions, the following results will be used frequently.

Basic Derivative Rules

Let $f : C \to \mathbb{R}$, with $C \subseteq \mathbb{R}$.

The derivative of f at x is defined as

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

We also use the notation

$$f'(x) = \frac{df(x)}{dx}.$$

$->$ The derivative is a local concept. It measures the rate of change of a function at a point.

Useful Derivatives

$$\frac{d}{dx}(c) = 0$$

$$\frac{d}{dx}(a + bx) = b$$

$$\frac{d}{dx}(ax^n) = anx^{n-1}$$

$$\frac{d}{dx}(e^{ax}) = ae^{ax}$$

$$\frac{d}{dx}(\ln x) = \frac{1}{x}$$

Derivative Rules for Combined Functions

Let f and g be differentiable functions.

$$\frac{d}{dx}(f(x) + g(x)) = f'(x) + g'(x)$$

$$\frac{d}{dx}(f(x) - g(x)) = f'(x) - g'(x)$$

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + g'(x)f(x)$$

If $g(x) \neq 0$,

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - g'(x)f(x)}{g(x)^2}$$

Chain rule:

$$\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$$

Concavity and Second Derivatives

Concavity plays a central role in optimization problems.

A function f is concave if

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

for all $x_1, x_2 \in C$ and $\lambda \in [0, 1]$.

If f is twice differentiable, then

$$f''(x) \leq 0 \quad \text{implies concavity.}$$

If $f''(x) < 0$, then f is strictly concave.

Local and Global Maxima

A point x^* is a local maximum if

$$f(x^*) \geq f(x) \quad \text{for all } x \text{ sufficiently close to } x^*.$$

It is a global maximum if

$$f(x^*) \geq f(x) \quad \text{for all } x \in C.$$

Strict concavity guarantees that a local maximum is automatically global.

Partial Derivatives

In many games, payoffs depend on several variables. For example,

$$\pi_i(q_i, q_j).$$

The partial derivative with respect to q_i is

$$\frac{\partial \pi_i}{\partial q_i}.$$

It measures how player i's payoff changes when q_i changes, holding other variables fixed.

Example: Two-Variable Function

Let

$$f(x, y) = 4x - x^2 - xy.$$

Then

$$\frac{\partial f}{\partial x} = 4 - 2x - y,$$

$$\frac{\partial f}{\partial y} = -x.$$

To find a critical point, we set both partial derivatives equal to zero.

Second partial derivatives help determine curvature in multivariable settings, which will be important when studying best responses and equilibrium.

1.4 Choices

Game theory begins with a simple idea: a decision maker must choose among alternatives, and each choice leads to consequences.

Basic Decision Theory

A decision maker (also called an agent or a player) faces a situation in which she must choose one of several alternatives.

Each choice leads to an outcome, and the consequences of that outcome are borne by the decision maker herself.

In order to analyze a decision problem, we must specify these three elements:

1. What actions are available?
2. What outcomes result from each action?
3. How does the decision maker rank those outcomes?

The Three Elements of a Decision Problem

A decision problem consists of three basic components:

- **Actions.** Actions are the set of alternatives from which the decision maker can choose. We denote the set of possible actions by A.
- **Outcomes.** Outcomes are the possible consequences that result from actions. We denote the set of outcomes by X.
- **Preferences.** Preferences describe how the decision maker ranks the set of possible outcomes, from most desired to least desired. Preferences are typically written using a binary relation $\succeq$.

The preference relation $x \succeq y$ means that outcome x is weakly preferred to outcome y.

We also define:

$$x \succ y \quad \text{if } x \succeq y \text{ and not } y \succeq x,$$

$$x \sim y \quad \text{if } x \succeq y \text{ and } y \succeq x.$$

Rational Preferences

To model rational decision making, we impose two fundamental properties on preferences.

- **Completeness.** For any two outcomes $x, y \in X$, the decision maker can compare them. That is,

  $$x \succeq y, \quad y \succeq x, \quad \text{or both.}$$

  Completeness means the decision maker is never unable to rank two alternatives.
- **Transitivity.** For any $x, y, z \in X$,

  $$x \succeq y \quad \text{and} \quad y \succeq z \quad \Rightarrow \quad x \succeq z.$$

  Transitivity guarantees internal consistency in ranking.

Preferences satisfying completeness and transitivity are called **rational preferences**.

Utility Representation

If preferences are rational, we can represent them numerically.

A function

$$u : X \to \mathbb{R}$$

represents preferences if for any $x, y \in X$,

$$u(x) \geq u(y) \quad \text{if and only if} \quad x \succeq y.$$

The function $u(x)$ is called a payoff function or utility function. Utility does not measure happiness in a psychological sense. It is simply a numerical representation of preferences.

Only the ranking matters. If u represents preferences, then any strictly increasing transformation of u represents the same preferences.

Example

Suppose a decision maker ranks outcomes as

$$x_1 \succ x_2 \succ x_3.$$

One possible utility representation is

$$u(x_1) = 3, \quad u(x_2) = 2, \quad u(x_3) = 1.$$

Another valid representation is

$$v(x_1) = 100, \quad v(x_2) = 50, \quad v(x_3) = 10.$$

Both represent the same preference ordering.

1.5 The Rational Choice Paradigm

The rational choice paradigm assumes that a decision maker chooses an action that maximizes her payoff (utility), given what she knows about the available actions and the consequences of those actions.

A decision problem is fully described once the decision maker knows:

1. all possible actions, A;
2. all possible outcomes, X;
3. how actions map into outcomes (what outcome happens under each action);
4. her rational preferences (payoffs) over outcomes.

Definition (Rational Choice):

Let A be the set of feasible actions and let $u(\cdot)$ be a payoff (utility) function defined over actions.

A decision maker is **rational** if she chooses an action $a^* \in A$ that maximizes her payoff:

$$a^* \in \arg\max_{a \in A} u(a).$$

Equivalently, $a^* \in A$ is chosen if and only if

$$u(a^*) \geq u(a) \quad \text{for all } a \in A.$$

Example 1

Action space: $A = [0, 1]$
Utility function:

$$u(a) = 2a - 4a^2.$$

Question: What is the rational choice?

We maximize $u(a)$ over $a \in [0, 1]$. First, we compute the derivative:

$$u'(a) = 2 - 8a.$$

Then set $u'(a) = 0$:

$$2 - 8a = 0 \quad \Rightarrow \quad a = \frac{1}{4}.$$

Check curvature:

$$u''(a) = -8 < 0,$$

so u is strictly concave and this critical point is the unique maximizer (as long as it is feasible). Since $\frac{1}{4} \in [0, 1]$, it is feasible. For completeness, check endpoints:

$$u(0) = 0, \qquad u(1) = 2 - 4 = -2.$$

Therefore, the rational choice is

$$\boxed{a^* = \frac{1}{4}.}$$

Example 2

Action space: $p \in \mathbb{R}_{\geq 0}$
Market demand:

$$q = 40 - 2p.$$

Question: What is the monopolist's rational choice (profit-maximizing price)?

With no cost information given, profit equals revenue:

$$\pi(p) = pq = p(40 - 2p) = 40p - 2p^2.$$

Maximize $\pi(p)$ over $p \geq 0$.
Derivative:

$$\pi'(p) = 40 - 4p.$$

Set $\pi'(p) = 0$:

$$40 - 4p = 0 \quad \Rightarrow \quad p = 10.$$

Second derivative:

$$\pi''(p) = -4 < 0,$$

so this is a maximum. So the profit-maximizing (rational) price is

$$\boxed{p^* = 10.}$$

The corresponding quantity is

$$q^* = 40 - 2(10) = 20.$$

Example 3

Action space: $A = \mathbb{R}_{\geq 0}$

Utility function:

$$u(a) = \ln(2a) - \frac{1}{4}a.$$

Question: What is the rational choice?

Because of the logarithm, we must have $a > 0$.

Differentiate:

$$u'(a) = \frac{d}{da}\ln(2a) - \frac{1}{4} = \frac{1}{a} - \frac{1}{4}.$$

Set $u'(a) = 0$:

$$\frac{1}{a} - \frac{1}{4} = 0 \quad \Rightarrow \quad \frac{1}{a} = \frac{1}{4} \quad \Rightarrow \quad a = 4.$$

Second derivative:

$$u''(a) = -\frac{1}{a^2} < 0 \quad \text{for } a > 0,$$

so u is strictly concave and $a = 4$ is the unique maximizer.

Therefore, the rational choice is

$$\boxed{a^* = 4.}$$

1.6 Uncertainty, Probability, and Expectations

So far, we analyzed decision problems without uncertainty. However, in many situations, the decision maker does not know with certainty which outcome will occur.

Uncertainty plays a central role in:

- individual decision problems (uncertainty about the state of the world),
- strategic games (uncertainty about opponents or about nature).

Here, we focus on stochastic (probabilistic) uncertainty and primarily on risk neutrality.

Probability Distributions

Let Ω denote a finite set of states of the world.

A probability distribution over Ω is a function

$$p \in \Delta(\Omega)$$

such that

$$p(\omega) \geq 0 \quad \forall \omega \in \Omega,$$

$$\sum_{\omega \in \Omega} p(\omega) = 1.$$

Thus, probabilities are non-negative and sum to one.

State-Dependent Utility

Under uncertainty, payoffs depend on both the chosen action and the realized state.

We define a state-dependent payoff function:

$$u : A \times \Omega \to \mathbb{R}.$$

For each action a and state ω, the realized payoff is $u(a, \omega)$.

Expected Value

Given a probability distribution p, the expected value of action a is

$$\mathbb{E}_p(a) = \sum_{\omega \in \Omega} p(\omega)u(a, \omega).$$

The expectation is the probability-weighted average payoff.

A risk-neutral decision maker chooses

$$a^* \in \arg\max_{a \in A} \mathbb{E}_p(a).$$

Risk neutrality therefore means maximizing expected value.

Example: Coin Flip

Suppose a decision maker chooses action a and receives

$$u(a, H) = 10, \qquad u(a, T) = 4.$$

If the coin is fair,

$$p(H) = \frac{1}{2}, \qquad p(T) = \frac{1}{2}.$$

The expected payoff is

$$\mathbb{E}_p(a) = \frac{1}{2}(10) + \frac{1}{2}(4) = 7.$$

If heads is three times as likely as tails, then

$$p(H) = \frac{3}{4}, \qquad p(T) = \frac{1}{4},$$

and

$$\mathbb{E}_p(a) = \frac{3}{4}(10) + \frac{1}{4}(4) = 8.5.$$

Example: Fair Die

Suppose a decision maker rolls a fair die and receives payoff equal to the number rolled:

$$u(a, \omega) = \omega, \qquad \omega \in \{1, \dots, 6\}.$$

Since the die is fair,

$$p(\omega) = \frac{1}{6}.$$

The expected payoff is

$$\mathbb{E}_p(a) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5.$$

Risk Attitudes

Risk attitudes describe how a decision maker evaluates uncertain payoffs.

- **Risk-neutral:** maximizes expected value.
- **Risk-averse:** prefers smoother, less extreme payoffs.
- **Risk-loving:** prefers more extreme payoffs.

Risk and the Shape of Utility

A decision maker's attitude toward risk is determined by the curvature of the utility function.

Theorem 1.1. *A decision maker is risk-averse if and only if the utility function u is concave.*

Theorem 1.2. *A decision maker is risk-loving if and only if the utility function u is convex.*

If $u''(x) < 0$, utility is concave and the decision maker dislikes risk. If $u''(x) > 0$, utility is convex and the decision maker likes risk.

Insurance Example

Suppose there is a 20% probability of incurring a \$100 loss.
The expected loss is

$$0.2 \times 100 = 20.$$

A risk-neutral decision maker would pay at most \$20 for full insurance.
A risk-averse decision maker would pay strictly more than \$20.

Lottery Example

Suppose a lottery ticket costs \$1 and pays \$1001 with probability p.
Expected value:

$$\mathbb{E} = 1001p - 1.$$

A risk-neutral agent buys the ticket if

$$1001p - 1 \geq 0 \quad \Rightarrow \quad p \geq \frac{1}{1001}.$$

A risk-loving agent may buy even if p is smaller.

Chapter 2: Normal Form Games

2.1 Strategic Games

A strategic game is a model of interaction between rational decision makers who choose their actions simultaneously. Each player has a set of possible actions, and the outcome of the game depends on the combination of all players' choices. The model specifies payoffs for every possible action profile, capturing each player's preferences over outcomes.

The Formal Definition

A **strategic game** consists of:

- A set of players.
- For each player, a set of actions.
- For each player, payoffs over the set of action profiles.

This definition applies to **static games**: players choose their actions **simultaneously**. We focus on **normal form games**.

Prisoner's Dilemma

- Two suspects in a major crime are held in separate cells.
- There is enough evidence to convict each of a minor offense.
- There is not enough evidence to convict either of the major crime unless one acts as an informer against the other (**Confess**).

Possible outcomes:

- If they both **Don't Confess**, each is convicted of a minor offense and spends 1 year in prison.
- If exactly one **Confesses**, the confessor is freed and used as a witness, while the other spends 3 years in prison.
- If they both **Confess**, each spends 2 years in prison.

Players, Actions, and Preferences

- Players: the two suspects.
- Actions: {Don't Confess, Confess}.

Define Don't Confess = D and Confess = C. Then suspect 1's payoffs can be ordered as

$$u_1(C,D) > u_1(D,D) > u_1(C,C) > u_1(D,C),$$

and suspect 2's payoffs can be ordered as

$$u_2(D,C) > u_2(D,D) > u_2(C,C) > u_2(C,D).$$

Figure: Payoff Matrix for the prisoner's dilemma.

		Player 2	
		C	D
Player 1	C	2 2	3 0
	D	0 3	1 1

Bach or Stravinsky (Marriage Game)

Two people wish to go out together. Two concerts are available: one of music by Bach (B), and one of music by Stravinsky (S). One person prefers Bach and the other prefers Stravinsky. If they go to different concerts, each of them is equally unhappy listening to the music of either composer.

Figure: Payoff Matrix for Bach or Stravinsky.

		Player 2	
		B	S
Player 1	B	1 2	0 0
	S	0 0	2 1

Matching Pennies

Two people choose simultaneously whether to show the Head (H) or the Tail (T) of a coin. If they show the same side, person 2 pays person 1 a dollar; if they show different sides, person 1 pays person 2 a dollar. Each person cares only about the amount of money she receives, and prefers to receive more than less.

Figure: Payoff Matrix for the matching pennies game.

		Player 2	
		H	T
Player 1	H	-1 1	1 -1
	T	1 -1	-1 1

2.2 Iterated Elimination of Strictly Dominated Strategies

Having described how to represent a strategic game, we now introduce a first method for solving a normal form game.

The guiding idea is simple: a rational player will never play a strictly dominated strategy.

Strict Dominance

Definition. Consider a normal form game with two players. Let $a_i', a_i'' \in A_i$.
The action (strategy) a_i' is *strictly dominated* by strategy a_i'' if for every action $a_j \in A_j$ of player j, player i's payoff from playing a_i' is strictly less than the payoff from playing a_i'':

$$u_i(a_i', a_j) < u_i(a_i'', a_j) \quad \forall a_j \in A_j.$$

In words, strategy a_i'' yields a strictly higher payoff than a_i' no matter what the other player does. A rational player therefore has no reason to choose a_i'.

Application: Prisoner's Dilemma

Recall the Prisoner's Dilemma with actions C (Confess) and D (Don't Confess).

For prisoner i:

- If the other prisoner plays D, then i prefers C to avoid prison.
- If the other prisoner plays C, then i prefers C to receive a shorter sentence.

Thus, for each strategy of player j, the payoff to player i from playing D is strictly less than the payoff from playing C.

Therefore, for each prisoner, strategy D is strictly dominated by strategy C.

Iterated Elimination

The procedure is:

1. Remove any strictly dominated strategies.
2. Reduce the game accordingly.
3. Check again whether new strategies become strictly dominated.
4. Repeat until no strictly dominated strategies remain.

This process is called *Iterated Elimination of Strictly Dominated Strategies*.

Limitations

Although the method is intuitive, it has important limitations.

- Each step requires an assumption about what players know about each other's rationality.
- To apply the process repeatedly, we must assume *common knowledge* of rationality.

That is, we assume:

- All players are rational.
- All players know that all players are rational.
- All players know that all players know that all players are rational.
- And so on, *ad infinitum*.

A further drawback is that iterated elimination often produces imprecise predictions about play. Moreover, it becomes increasingly complex in games with more than two players.

2.3 Introducing Nash Equilibrium

What actions will be chosen by the players in a strategic game?

To answer this question, we need an equilibrium concept. In this course we use the notion of **Nash equilibrium** (and its extensions).

Intuition

Intuitively, a Nash equilibrium has two components:

1. Each player chooses her action according to rational choice, given her belief about the other players' actions.
2. Each player's belief about the other players' actions is correct.

Thus, in equilibrium, every player is best responding to the actions actually chosen by the others.

Informal Description

A Nash equilibrium is an action profile a^* with the property that no player can do better by unilaterally deviating, given that all other players adhere to a^*.

Formal Definition

Definition (Nash Equilibrium):
Let $a^* = (a_1^*, \ldots, a_n^*)$ be an action profile in a strategic game.
The action profile a^* is a Nash equilibrium if, for every player i and every action $a_i \in A_i$,

$$u_i(a_i^*, a_{-i}^*) \;\geq\; u_i(a_i, a_{-i}^*).$$

That is, given that every other player $j \neq i$ chooses a_j^*, player i cannot obtain a higher payoff by choosing any alternative action a_i.
Here u_i is the payoff function representing player i's preferences.

2.4 Solving Games

We now apply best response functions to solve games and identify Nash equilibria.

Bach or Stravinsky (Marriage Game)

Recall the payoff matrix:

Figure: Payoff Matrix for Bach or Stravinsky.

		Player 2	
		B	S
Player 1	B	1 2	0 0
	S	0 0	2 1

Best Responses: Player 1

- If Player 2 plays B, Player 1 prefers B since $2 > 0$.
- If Player 2 plays S, Player 1 prefers S since $1 > 0$.

Thus,

$$B_1(B) = \{B\}, \qquad B_1(S) = \{S\}.$$

Best Responses:Player 2

- If Player 1 plays B, Player 2 prefers B since $1 > 0$.
- If Player 1 plays S, Player 2 prefers S since $2 > 0$.

So,

$$B_2(B) = \{B\}, \qquad B_2(S) = \{S\}.$$

Nash Equilibria

The Nash equilibria are the action profiles where both players are playing best responses simultaneously.

Hence, there are two pure-strategy Nash equilibria:

$$(B, B) \quad \text{and} \quad (S, S).$$

This game has multiple equilibria.

Matching Pennies

Recall the payoff matrix:

Figure: Payoff Matrix for Matching Pennies.

		Player 2	
		H	*T*
Player 1	*H*	-1 1	1 -1
	T	1 -1	-1 1

Best Responses: Player 1

- If Player 2 plays H, Player 1 prefers H.
- If Player 2 plays T, Player 1 prefers T.

Best Responses: Player 2

- If Player 1 plays H, Player 2 prefers T.
- If Player 1 plays T, Player 2 prefers H.

Nash Equilibrium

There is no action profile in which both players are simultaneously best responding. Therefore, **Matching Pennies has no pure-strategy Nash equilibrium.**

2.5 Best Response and Nash Equilibria

The concept of best responses allows us to systematically identify Nash equilibria in strategic games.

Best Response Functions

Definition (Best Response Function).
In an n-player strategic game, the best response function of player i is

$$B_i : \prod_{j \neq i} A_j \to A_i,$$

where for every action profile of the other players a_{-i},

$$B_i(a_{-i}) = \left\{ a_i \in A_i \ : \ u_i(a_i, a_{-i}) \geq u_i(a_i', a_{-i}) \text{ for all } a_i' \in A_i \right\}.$$

What it means.

The best response function tells us what player i should choose given what everyone else is doing. It lists the actions that maximize player i's payoff, holding the other players' actions fixed.

In a two-player game this simplifies to

$$B_1 : A_2 \to A_1 \qquad \text{and} \qquad B_2 : A_1 \to A_2.$$

Each player's best response depends only on the other player's action.

Nash Equilibrium

Definition (Nash Equilibrium).
An action profile $a^* = (a_1^*, \ldots, a_n^*)$ is a Nash equilibrium if

$$a_i^* \in B_i(a_{-i}^*) \quad \text{for every player } i.$$

What it means.

A Nash equilibrium is a situation where every player is best responding to the others. No one has an incentive to deviate unilaterally. If any single player changes their action while others keep theirs fixed, their payoff does not increase.

Strict Nash Equilibrium

Definition (Strict Nash Equilibrium).
An action profile a^* is a strict Nash equilibrium if for every player i,

$$u_i(a^*) > u_i(a_i, a^*_{-i}) \quad \text{for every } a_i \neq a^*_i.$$

What it means.

In a strict Nash equilibrium, each player's equilibrium action is strictly better than any other action. No ties are allowed. This makes the equilibrium more robust.

Pareto Dominance and Pareto Optimality

Definition (Pareto Dominance).
An action profile a^* Pareto dominates another profile a' if

$$u_i(a^*) \geq u_i(a') \quad \text{for all players } i,$$

and

$$u_j(a^*) > u_j(a') \quad \text{for at least one player } j.$$

What it means.

One outcome Pareto dominates another if everyone is at least as well off and someone is strictly better off. It represents an improvement that harms no one.

Definition (Pareto Optimality).
An action profile is Pareto optimal if it is not Pareto dominated by any other action profile.

What it means.

A Pareto optimal outcome is efficient. There is no way to make someone better off without making someone else worse off.

Dominated Strategies

Definition (Strict Dominance).
An action a_i'' strictly dominates a_i' if

$$u_i(a_i'', a_{-i}) > u_i(a_i', a_{-i}) \quad \text{for every } a_{-i}.$$

What it means.

A strictly dominated strategy is never optimal. It gives a lower payoff no matter what others do, so a rational player should never choose it.

Definition (Weak Dominance).
An action a_i'' weakly dominates a_i' if

$$u_i(a_i'', a_{-i}) \geq u_i(a_i', a_{-i}) \quad \text{for every } a_{-i},$$

and

$$u_i(a_i'', a_{-i}) > u_i(a_i', a_{-i}) \quad \text{for some } a_{-i}.$$

What it means.

A weakly dominated strategy is never better and sometimes strictly worse. It might tie in some cases, but there exists at least one situation where it performs strictly worse.

Examples

Example 1 (Two-player game).

		Player 2		
		X	Y	Z
Player 1	A	$(4,0)$	$(2,5)$	$(1,3)$
	B	$(7,6)$	$(0,4)$	$(2,2)$
	C	$(3,1)$	$(5,2)$	$(1,4)$

1) Best response functions.

For Player 1 (rows), compare Player 1's payoff across rows in each column:

$$B_1(X) = \{B\}, \qquad B_1(Y) = \{C\}, \qquad B_1(Z) = \{B\}.$$

For Player 2 (columns), compare Player 2's payoff across columns in each row:

$$B_2(A) = \{Y\}, \qquad B_2(B) = \{X\}, \qquad B_2(C) = \{Z\}.$$

2) Nash equilibrium.

A pure-strategy Nash equilibrium is a cell where each action is a best response to the other. The only mutual best response is:

$$(B, X),$$

with payoff $(7,6)$.

3) Pareto optimality.

(B, X) is Pareto optimal. Player 2's payoff is 6, and there is no other outcome that gives Player 2 at least 6 while also making Player 1 at least 7 (or even keeping Player 1 at least as well off). So no other action profile Pareto dominates (B, X).

Example 2 (Three-player game).
Player 1 chooses U or D. Player 2 chooses L or R. Player 3 chooses A or B.
The payoff matrices below show (u_1, u_2, u_3). The left matrix is when Player 3 plays A, and the right matrix is when Player 3 plays B.

A

		Player 2	
		L	R
Player 1	U	$(3, 2, 3)$	$(1, 2, 0)$
	D	$(1, 2, 0)$	$(1, 2, 0)$

B

		Player 2	
		L	R
Player 1	U	$(1, 4, 2)$	$(1, 2, 0)$
	D	$(1, 2, 0)$	$(3, 4, 1)$

1) Best response functions.
Player 1: best response to (a_2, a_3).

$$B_1(L, A) = \{U\}, \quad B_1(R, A) = \{U, D\}, \quad B_1(L, B) = \{U, D\}, \quad B_1(R, B) = \{D\}.$$

Player 2: best response to (a_1, a_3).

$$B_2(U, A) = \{L, R\}, \quad B_2(D, A) = \{L, R\}, \quad B_2(U, B) = \{L\}, \quad B_2(D, B) = \{R\}.$$

Player 3: best response to (a_1, a_2).

$$B_3(U, L) = \{A\}, \quad B_3(U, R) = \{A, B\}, \quad B_3(D, L) = \{A, B\}, \quad B_3(D, R) = \{B\}.$$

2) Nash equilibria: A profile (a_1, a_2, a_3) is a Nash equilibrium if each $a_i \in B_i(a_{-i})$. The pure-strategy Nash equilibria are:

$$(U, L, A), \qquad (U, R, A), \qquad (D, R, B).$$

Their payoffs are: $(U, L, A) : (3, 2, 3)$, $(U, R, A) : (1, 2, 0)$, $(D, R, B) : (3, 4, 1)$.
3) Pareto optimality: (U, R, A) is *not* Pareto optimal because (U, L, A) yields $(3, 2, 3) \geq (1, 2, 0)$ componentwise, with strict improvements for Players 1 and 3.
(U, L, A) *is* Pareto optimal (no other profile makes all three weakly better and at least one strictly better).
(D, R, B) *is* Pareto optimal (there is no other profile that weakly increases all three payoffs relative to $(3, 4, 1)$ while strictly improving at least one).

Symmetric Games

Def (Symmetric game). A two-player strategic game is *symmetric* if the players' sets of actions are the same and the players' preferences are represented by payoff functions u_1 and u_2 such that

$$u_1(a_1, a_2) = u_2(a_2, a_1) \quad \text{for every pair of actions } (a_1, a_2).$$

This means the game treats the two players identically. If we swap the players' roles, the payoffs swap in the same way.

Def (Symmetric Nash equilibrium). In a strategic game in which each player has the same set of actions, an action profile a^* is a *symmetric Nash equilibrium* if it is a Nash equilibrium and

$$a_i^* = a_j^* \quad \text{for all players } i, j.$$

So a symmetric Nash equilibrium is a Nash equilibrium where everyone plays the same action. A symmetric game can still have *no* symmetric Nash equilibrium.

Example (A symmetric game with many Nash equilibria)

Let $S = [0, 1]$ and define for each player i:

$$u_i(s_1, s_2) = \begin{cases} \max\{s_1, s_2\} & \text{if } (s_1, s_2) \neq (1, 1), \\ 0 & \text{if } (s_1, s_2) = (1, 1). \end{cases}$$

Claim: The Nash equilibrium set is

$$\left\{(s_1, s_2) \in [0, 1]^2 : s_1 = 1 \text{ or } s_2 = 1\right\} \setminus \{(1, 1)\}.$$

If $s_2 < 1$, player 1 can improve by choosing $s_1 = 1$, since then $\max\{s_1, s_2\} = 1$ and the outcome is not $(1, 1)$. So in any Nash equilibrium with $s_2 < 1$, we must have $s_1 = 1$. Symmetrically, if $s_1 < 1$, we must have $s_2 = 1$. Finally, $(1, 1)$ is not a Nash equilibrium because it gives payoff 0, but a player can deviate to (say) $1 - \varepsilon$ and get payoff 1.

Chapter 3: Applications of Nash Equilibrium

3.1 Cournot's Oligopoly

Model Setup

Consider an industry with two firms competing in quantities (Cournot competition).

- Firms are symmetric.
- Constant marginal cost: $c > 0$.
- No fixed costs.
- Linear inverse demand:

  $$p(Q) = a - bQ,$$

  where $a > c$, $b > 0$, and

  $$Q = q_i + q_j$$

  is total industry output.

The condition $a > c$ ensures that producing a positive quantity is profitable, since the highest willingness to pay exceeds marginal cost.

Profit Maximization Problem

Firm i's profit is:

$$\pi_i(q_i, q_j) = p(Q)q_i - cq_i$$

Substituting demand:

$$\pi_i(q_i, q_j) = (a - b(q_i + q_j))q_i - cq_i$$

Each firm chooses $q_i \geq 0$, taking q_j as given:

$$\max_{q_i \geq 0} \quad (a - bq_i - bq_j)q_i - cq_i$$

Best Response Function

Take the first-order condition with respect to q_i:

$$\frac{\partial \pi_i}{\partial q_i} = a - 2bq_i - bq_j - c = 0$$

Solving for q_i:

$$BR_i(q_j) = q_i = \frac{a - c - bq_j}{2b}$$

By symmetry:

$$BR_j(q_i) = \frac{a - c - bq_i}{2b}$$

Each firm's best response is downward sloping in the rival's output:

$$\frac{\partial q_i}{\partial q_j} = -\frac{1}{2}$$

Thus quantities are **strategic substitutes**.

Nash Equilibrium

The Nash equilibrium is where best response functions intersect.

Solve:

$$q_i = \frac{a - c - bq_j}{2b}$$

$$q_j = \frac{a - c - bq_i}{2b}$$

By symmetry, let $q_i = q_j = q^*$:

$$q^* = \frac{a - c - bq^*}{2b}$$

Multiply both sides by $2b$:

$$2bq^* = a - c - bq^*$$

$$3bq^* = a - c$$

$$\boxed{q_i^* = q_j^* = \frac{a-c}{3b}}$$

Total output:

$$Q^* = \frac{2(a-c)}{3b}$$

Price:

$$p^* = a - bQ^* = a - \frac{2(a-c)}{3} = \frac{a+2c}{3}$$

Graphical Representation

The Nash equilibrium is the intersection of the two best response functions.

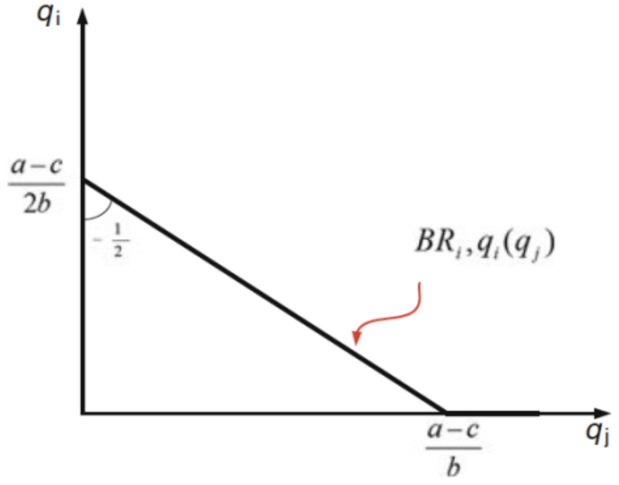


Figure 3.1: Best response functions and Nash equilibrium.

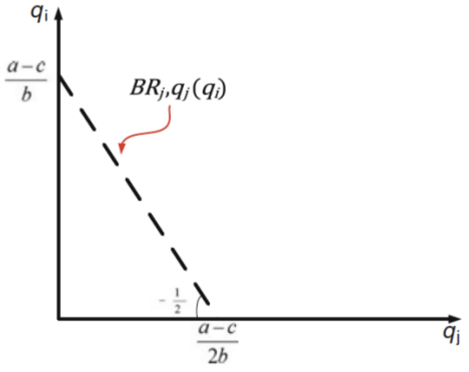


Figure 3.2: Best response of firm j.

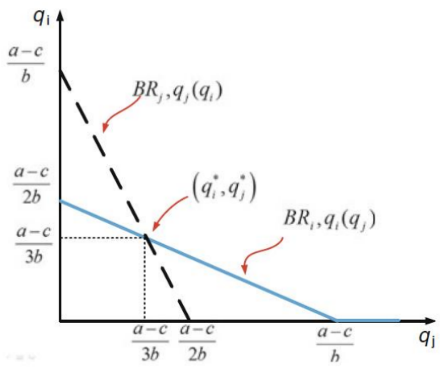


Figure 3.3: Best response of firm i.

The intersection point represents a Nash equilibrium because:

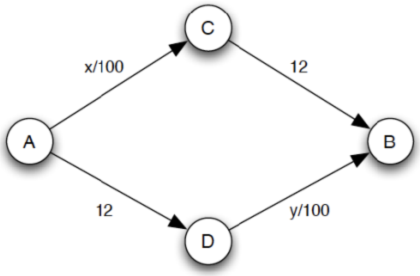
- Given q_j^*, firm i does not want to deviate.
- Given q_i^*, firm j does not want to deviate.

3.2 Examples

Congestion Game

Setup.

There are 1000 drivers traveling from A to B.



There are two possible routes:

- Upper route: $A \to C \to B$
- Lower route: $A \to D \to B$

Let:

$$x = \text{number of cars using edge } AC, \quad y = \text{number of cars using edge } DB.$$

Travel times:

$$\text{Time on } AC = \frac{x}{100}, \quad \text{Time on } DB = \frac{y}{100}.$$

$$\text{Time on } CB = 12, \quad \text{Time on } AD = 12.$$

Since all cars must travel from A to B,

$$x + y = 1000.$$

Each driver chooses a route to minimize travel time.

1) Nash equilibrium without the new road

Upper route time:

$$T_U = \frac{x}{100} + 12.$$

Lower route time:

$$T_L = 12 + \frac{y}{100}.$$

In equilibrium, any used routes must have equal travel times. Thus,

$$\frac{x}{100} + 12 = 12 + \frac{y}{100}.$$

Cancel 12:

$$\frac{x}{100} = \frac{y}{100}.$$

Thus:

$$x = y.$$

Since $x + y = 1000$,

$$x = y = 500.$$

Equilibrium travel time:

$$T = \frac{500}{100} + 12 = 5 + 12 = 17.$$

Total travel cost:

$$1000 \times 17 = 17{,}000.$$

2) After adding the new road $C \to D$ (zero cost)

Now a new route exists:

$$A \to C \to D \to B.$$

Its travel time is:

$$T_{ACDB} = \frac{x}{100} + 0 + \frac{y}{100}.$$

Now drivers compare:

$$T_{ACB} = \frac{x}{100} + 12,$$

$$T_{ADB} = 12 + \frac{y}{100},$$

$$T_{ACDB} = \frac{x}{100} + \frac{y}{100}.$$

In equilibrium, no driver can improve by switching routes.

Suppose all drivers use the new route. Then:

$$x = 1000, \quad y = 1000.$$

Then:

$$T_{ACDB} = \frac{1000}{100} + \frac{1000}{100} = 10 + 10 = 20.$$

Check deviations:

Upper route:

$$\frac{1000}{100} + 12 = 10 + 12 = 22.$$

Lower route:

$$12 + \frac{1000}{100} = 12 + 10 = 22.$$

Thus, 20 is strictly lower than 22, so no one deviates.

New Nash equilibrium:

$$x = 1000, \quad y = 1000.$$

Equilibrium travel time:

$$T = 20.$$

Total travel cost:

$$1000 \times 20 = 20{,}000.$$

Conclusion (Braess Paradox).

Adding the zero-cost road increases total travel time:

$$17{,}000 \quad \longrightarrow \quad 20{,}000.$$

The network becomes strictly worse after adding a new road.

War of Attrition

The War of Attrition is a strategic game in which two players compete over an indivisible object. Each player chooses a *concession time*, and waiting is costly.

Model

- **Players:** $i = 1, 2$.
- **Strategies:** Each player chooses a concession time $t_i \in [0, \infty)$.
- **Valuations:** Player i values the object at $v_i > 0$.
- **Cost of waiting:** Each unit of time costs 1.

Payoffs are given by

$$u_i(t_i, t_j) = \begin{cases} -t_i & \text{if } t_i < t_j, \\ \frac{1}{2}v_i - t_i & \text{if } t_i = t_j, \\ v_i - t_j & \text{if } t_i > t_j, \end{cases}$$

where $j \neq i$.

If player i concedes first, she pays the waiting cost and receives nothing. If the opponent concedes first, player i obtains the object but still incurs the waiting cost. If both concede at the same time, they split the object equally.

Best Responses

Fix t_j. Player i's optimal choice depends on the comparison between t_j and v_i:

- If $t_j < v_i$, it is optimal to wait longer than t_j.
- If $t_j = v_i$, player i is indifferent between conceding immediately and waiting longer than t_j.
- If $t_j > v_i$, it is optimal to concede immediately.

The best response correspondence is therefore

$$B_i(t_j) = \begin{cases} \{t_i : t_i > t_j\} & \text{if } t_j < v_i, \\ \{0\} \cup \{t_i : t_i > t_j\} & \text{if } t_j = v_i, \\ \{0\} & \text{if } t_j > v_i. \end{cases}$$

Nash Equilibria

A profile (t_1, t_2) is a Nash equilibrium if and only if either

$$t_1 = 0 \quad \text{and} \quad t_2 \geq v_1,$$

or

$$t_2 = 0 \quad \text{and} \quad t_1 \geq v_2.$$

In every equilibrium, one player concedes immediately and the other waits at least as long as the opponent's valuation. There is no symmetric pure strategy equilibrium.

Tournaments

Many strategic environments can be modeled as tournaments, where the probability of winning a prize depends on relative effort rather than absolute performance.

Model

Two firms compete in an R&D race.

- **Prize:** The winner receives 36.
- **Strategies:** Each firm chooses expenditure $x_i \in [0, 25]$.
- **Winning probability:**
$$\Pr(i \text{ wins}) = \frac{x_i}{x_1 + x_2}.$$
- **Cost:** Expenditure equals cost.

Firm 1's profit is

$$\pi_1(x_1, x_2) = 36\frac{x_1}{x_1 + x_2} - x_1,$$

and symmetrically,

$$\pi_2(x_1, x_2) = 36\frac{x_2}{x_1 + x_2} - x_2.$$

Best Response Functions

Firm 1 maximizes π_1 with respect to x_1. The first order condition is

$$\frac{\partial \pi_1}{\partial x_1} = 36\frac{x_2}{(x_1 + x_2)^2} - 1 = 0.$$

Rearranging,

$$36x_2 = (x_1 + x_2)^2.$$

Taking square roots,

$$6\sqrt{x_2} = x_1 + x_2.$$

Solving for x_1 gives firm 1's best response:

$$B_1(x_2) = 6\sqrt{x_2} - x_2.$$

By symmetry, firm 2's best response is

$$B_2(x_1) = 6\sqrt{x_1} - x_1.$$

Symmetric Nash Equilibrium

In a symmetric Nash equilibrium $x_1 = x_2 = x^*$. Substituting into the best response function:

$$x^* = 6\sqrt{x^*} - x^*.$$

Rearranging,

$$2x^* = 6\sqrt{x^*}.$$

Squaring both sides,

$$4(x^*)^2 = 36x^*,$$

which implies

$$x^* = 9.$$

Equilibrium Outcome

In equilibrium, both firms spend

$$x_1^* = x_2^* = 9.$$

Each firm wins with probability $\frac{1}{2}$ and earns expected profit

$$\pi_i = 36\frac{9}{18} - 9 = 18 - 9 = 9.$$

Chapter 4: Mixed Strategies Nash Equilibrium

4.1 Why Pure Strategies Are Not Always Enough

So far, whenever we found a Nash equilibrium, each player chose one specific action with certainty — "I will play H," or "I will play B." This is called a **pure strategy**. But recall what happened with Matching Pennies: no matter what action profile we looked at, one of the two players always wanted to deviate. There was simply no pure strategy Nash equilibrium.

This is not just a quirk of that one game. It turns out that in many strategic situations, **any predictable action can be exploited**. Think about penalty kicks in soccer: if a goalkeeper always dives left, the kicker will always shoot right. The goalkeeper's predictability is their weakness. The natural defense is to be **unpredictable**, to randomize.

So we need a richer notion of strategy, one that allows players to assign *probabilities* to their actions rather than committing to just one. This is the idea of a **mixed strategy**.

4.2 von Neumann and Morgenstern Payoff Functions

The Basic Idea

When players randomize, the outcome of the game is no longer certain, it becomes a **lottery** over action profiles. To evaluate lotteries and compare them rationally, we need a way to represent preferences over uncertain outcomes.

The standard approach, developed by von Neumann and Morgenstern (1944), gives us exactly this. Suppose a decision-maker has preferences over a set of outcomes, and each of her actions results in a probability distribution (lottery) over those outcomes. The vNM approach tells us that (under reasonable assumptions on those preferences) we can find a payoff function u over *deterministic* outcomes such that the decision-maker's preferences over lotteries are fully captured by the **expected value** of u.

vNM Expected Payoff Representation

Let there be K possible deterministic outcomes $a_1, a_2, \dots, a_K$, and let $(p_1, \dots, p_K)$ be a lottery where outcome a_k occurs with probability p_k.
Then the vNM payoff (expected utility) of this lottery is:

$$U(p_1, \dots, p_K) = \sum_{k=1}^{K} p_k \, u(a_k),$$

where $u(a_k)$ is the Bernoulli payoff assigned to outcome a_k.
A decision-maker prefers lottery $(p_1, \dots, p_K)$ over lottery $(p'_1, \dots, p'_K)$ if and only if:

$$\sum_{k=1}^{K} p_k \, u(a_k) > \sum_{k=1}^{K} p'_k \, u(a_k).$$

In plain language: to compare two lotteries, compute the probability-weighted average payoff of each, and pick the one with the higher expected payoff.

The function u that appears in this formula is called the **Bernoulli utility function**. It assigns a number to each deterministic outcome. Once we have u, evaluating any lottery is straightforward.

Example: Evaluating Two Lotteries

Suppose there are three outcomes: Win, Tie, Lose. A player's Bernoulli payoffs are:

$$u(\text{Win}) = 10, \qquad u(\text{Tie}) = 4, \qquad u(\text{Lose}) = 0.$$

Lottery A: Win with probability $\frac{1}{2}$, Lose with probability $\frac{1}{2}$.

$$U_A = \frac{1}{2}(10) + \frac{1}{2}(0) = 5.$$

Lottery B: Win with probability $\frac{1}{4}$, Tie with probability $\frac{3}{4}$.

$$U_B = \frac{1}{4}(10) + \frac{3}{4}(4) = 2.5 + 3 = 5.5.$$

Since $U_B > U_A$, the player prefers Lottery B.

The Simplex: Describing Mixed Strategies Formally

A mixed strategy for player i is simply a **probability distribution over her actions** $A_i = \{a_1, \dots, a_K\}$.

We denote the set of all possible mixed strategies for player i as Δ_i, and call it the **simplex** over A_i:

$$\Delta_i = \{\alpha_i \ : \ \alpha_i(a_1) + \cdots + \alpha_i(a_K) = 1, \ \ \alpha_i(a_k) \geq 0, \ \ k = 1, \dots, K\}.$$

Here $\alpha_i(a_k)$ is the probability that player i plays action a_k. The two conditions just say that probabilities must be non-negative and must sum to one, exactly what you'd expect.

A **pure strategy** is just a special mixed strategy where one action gets probability 1 and all others get probability 0. So pure strategies are contained inside the simplex, they are the "corners" of Δ_i. Mixed strategies are everything in between.

For example, if player i has two actions $\{H, T\}$, then a mixed strategy is fully described by a single number $p \in [0, 1]$:

$$\alpha_i(H) = p, \qquad \alpha_i(T) = 1 - p.$$

Playing $p = 1$ means always playing H (a pure strategy). Playing $p = 0$ means always playing T. Anything in between is a genuinely mixed strategy.

4.3 Beliefs and Mixed Strategies

What Is a Belief?

When players randomize, each player faces uncertainty about what the others will do. To analyze this, we introduce the concept of a **belief**.

Definition: Belief

A **belief** for player i is a probability distribution over the strategies of her opponents.

In a two-player game, player 1's belief is simply a probability distribution over player 2's actions — that is, player 1's best guess about what player 2 will do. In a Mixed Strategy Nash Equilibrium, these beliefs will turn out to be *correct*: each player's belief about the other's strategy matches what the other is actually playing.

Computing Expected Payoffs Under Mixed Strategies

Once both players randomize, payoffs are no longer certain numbers. They become expected payoffs. Here is how to compute them.

Suppose player 1 plays action a_1 with certainty, and player 2 mixes, playing action L with probability q and action R with probability $1 - q$.

Player 1's expected payoff from playing a_1 is:

$$U_1(a_1) = q \cdot u_1(a_1, L) + (1 - q) \cdot u_1(a_1, R).$$

Now suppose player 1 also mixes — playing a_1 with probability p and a'_1 with probability $1 - p$. Then:

$$U_1 = p \cdot [q \cdot u_1(a_1, L) + (1 - q) \cdot u_1(a_1, R)] + (1 - p) \cdot \left[q \cdot u_1(a'_1, L) + (1 - q) \cdot u_1(a'_1, R)\right].$$

Key Observation: Linearity

Expected payoffs are **linear** in each player's own mixing probability. This is a crucial property. It means that if playing H gives a higher expected payoff than playing T, then a player should put all probability on H (not mix). A player will only genuinely want to randomize when both actions give the *same* expected payoff.

4.4 Nash Equilibrium and Mixed Strategies

A strategic game with vNM preferences consists of a set of players, for each player a set of mixed strategies Δ_i, and for each player preferences over lotteries represented by the expected value of a payoff function over action profiles.

A Mixed Strategy Nash Equilibrium (MSNE) has two requirements:

1. Each player chooses her mixed strategy rationally, given her belief about what others will do.
2. Every player's belief about the others' strategies is correct.

Definition: Mixed Strategy Nash Equilibrium

A mixed strategy profile α^* is a Mixed Strategy Nash Equilibrium if, for each player i and every mixed strategy α_i,

$$U_i(\alpha^*) = U_i(\alpha_i^*, \alpha_{-i}^*) \geq U_i(\alpha_i, \alpha_{-i}^*),$$

where $U_i(\alpha)$ is player i's expected payoff to the mixed strategy profile α.

This is exactly the same idea as a pure strategy NE, just extended to allow randomization. Each player's mixed strategy must be a best response to everyone else's mixed strategy.

4.5 Best Response Functions

Definition: Best Response in Mixed Strategies

For a given α_{-i}, player i's best response function is:

$$B_i(\alpha_{-i}) = \{\alpha_i \in \Delta_i : U_i(\alpha_i, \alpha_{-i}) \geq U_i(\alpha_i', \alpha_{-i}) \text{ for all } \alpha_i' \in \Delta_i\}.$$

The mixed strategy profile α^* is a MSNE if and only if $\alpha_i^* \in B_i(\alpha_{-i}^*)$ for every player i.

Two-Player Two-Action Games

In a two-player two-action game, the setup: Player 1 chooses U or D, and Player 2 chooses L or R. We use shorthand notation:

$$p = \alpha_1(U), \quad 1 - p = \alpha_1(D), \qquad q = \alpha_2(L), \quad 1 - q = \alpha_2(R).$$

Since the players' choices are **independent**, the probability of any action pair (a_1, a_2) is the product of each player's probability of choosing that action:

$$\Pr(U, L) = p \cdot q, \quad \Pr(U, R) = p(1 - q), \quad \Pr(D, L) = (1 - p)q, \quad \Pr(D, R) = (1 - p)(1 - q).$$

Player 1's expected payoff is then:

$$U_1(p, q) = pq \cdot u_1(U, L) + p(1 - q) \cdot u_1(U, R) + (1 - p)q \cdot u_1(D, L) + (1 - p)(1 - q) \cdot u_1(D, R).$$

Because U_1 is **linear** in p, player 1 will want to put all weight on U if U gives a higher expected payoff than D, all weight on D if D is better, and will be willing to mix only when the two actions give *equal* expected payoffs.

The Indifference Principle

To find a mixed strategy Nash equilibrium in a 2×2 game:

- Choose Player 2's probability q so that Player 1 is **indifferent** between U and D.
- Choose Player 1's probability p so that Player 2 is **indifferent** between L and R.

Each player's mixing probability is determined by the *other* player's payoffs, not their own.

4.6 Applications

Matching Pennies

The payoff matrix:

		Player 2	
		H	T
Player 1	H	-1 1	1 -1
	T	1 -1	-1 1

We already showed this game has no pure strategy NE. Let $p = \Pr(\text{Player 1 plays } H)$ and $q = \Pr(\text{Player 2 plays } H)$.

Step 1: Find q so that Player 1 is indifferent between H and T.

Player 1's expected payoff from H:

$$U_1(H) = q(1) + (1-q)(-1) = q - 1 + q = 2q - 1.$$

Player 1's expected payoff from T:

$$U_1(T) = q(-1) + (1-q)(1) = -q + 1 - q = 1 - 2q.$$

Set $U_1(H) = U_1(T)$:

$$2q - 1 = 1 - 2q \implies 4q = 2 \implies q = \frac{1}{2}.$$

Step 2: Find p so that Player 2 is indifferent between H and T.

Player 2's expected payoff from H:

$$U_2(H) = p(-1) + (1 - p)(1) = -p + 1 - p = 1 - 2p.$$

Player 2's expected payoff from T:

$$U_2(T) = p(1) + (1 - p)(-1) = p - 1 + p = 2p - 1.$$

Set $U_2(H) = U_2(T)$:

$$1 - 2p = 2p - 1 \implies 2 = 4p \implies p = \frac{1}{2}.$$

MSNE: Both players play H and T each with probability $\frac{1}{2}$.

Verification: When $q = \frac{1}{2}$, Player 1 gets $U_1(H) = 2(\frac{1}{2}) - 1 = 0$ and $U_1(T) = 1 - 2(\frac{1}{2}) = 0$. Player 1 is indeed indifferent and willing to mix. The same holds for Player 2 when $p = \frac{1}{2}$. Each player's expected payoff in equilibrium is 0.

Bach or Stravinsky

Recall the payoff matrix, where Player 1 prefers Bach (B) and Player 2 prefers Stravinsky (S):

		Player 2	
		B	S
Player 1	B	2, 1	0, 0
	S	0, 0	1, 2

We already know there are two pure strategy NE: (B, B) and (S, S). Now we find the mixed strategy NE.

Let $p = \Pr(\text{Player 1 plays } B)$ and $q = \Pr(\text{Player 2 plays } B)$.

Step 1: Find q so that Player 1 is indifferent between B and S.

$$U_1(B) = q(2) + (1-q)(0) = 2q.$$

$$U_1(S) = q(0) + (1-q)(1) = 1-q.$$

Set $U_1(B) = U_1(S)$:

$$2q = 1-q \implies 3q = 1 \implies q = \frac{1}{3}.$$

Step 2: Find p so that Player 2 is indifferent between B and S.

$$U_2(B) = p(1) + (1-p)(0) = p.$$

$$U_2(S) = p(0) + (1-p)(2) = 2-2p.$$

Set $U_2(B) = U_2(S)$:

$$p = 2-2p \implies 3p = 2 \implies p = \frac{2}{3}.$$

MSNE: Player 1 plays B with probability $\frac{2}{3}$ and S with probability $\frac{1}{3}$. Player 2 plays B with probability $\frac{1}{3}$ and S with probability $\frac{2}{3}$.

Expected payoffs in equilibrium:

$$U_1 = 2q = 2 \cdot \frac{1}{3} = \frac{2}{3}, \qquad U_2 = p = \frac{2}{3}.$$

Note that both players earn $\frac{2}{3}$ in the mixed NE, which is lower than what they each get in the pure strategy NE they prefer (which gives payoff 2 to their favorite and 1 to the other). The mixed equilibrium is the least preferred of the three.

Coordination Game

Two players can accomplish a task only if both exert effort. The cost of effort is c, where $0 < c < 1$.

		Player 2	
		No Effort	Effort
Player 1	No Effort	0 0	$-c$ 0
	Effort	0 $-c$	$1-c$ $1-c$

There are two pure strategy NE: (No Effort, No Effort) and (Effort, Effort). Let $p = \Pr(\text{Player 1 exerts Effort})$ and $q = \Pr(\text{Player 2 exerts Effort})$.

Step 1: Find q so that Player 1 is indifferent.

$$U_1(\text{No Effort}) = q(0) + (1-q)(0) = 0.$$

$$U_1(\text{Effort}) = q(1-c) + (1-q)(-c) = q - qc - c + qc = q - c.$$

Set $U_1(\text{No Effort}) = U_1(\text{Effort})$:

$$0 = q - c \implies q = c.$$

Step 2: By symmetry, $p = c$.

MSNE: Each player exerts effort with probability c and shirks with probability $1 - c$.

Intuition: the higher the cost c, the more likely each player is to exert effort in the mixed NE. When effort is very costly, each player needs to believe the other is quite likely to exert effort before they are willing to do so themselves.

4.7 Characterizing Mixed Strategies

A player's expected payoff under a mixed strategy profile α is a weighted average of her expected payoffs to each pure action, where the weight on action a_i is the probability $\alpha_i(a_i)$:

$$U_i(\alpha) = \sum_{a_i \in A_i} \alpha_i(a_i)\, U_i(a_i, \alpha_{-i}).$$

This linearity leads to a very useful characterization of MSNE:

Proposition: Characterizing MSNE

A mixed strategy profile α^* is a MSNE if and only if, for each player i:

1. Every action played with positive probability yields the same expected payoff.
2. Every action played with zero probability yields an expected payoff *no greater than* the equilibrium payoff.

This is why we use the indifference condition to solve for MSNE: if a player genuinely mixes between two actions, she must be indifferent between them. If one action were strictly better, she would put all weight on it.

4.8 Existence of Nash Equilibrium

Theorem 4.1 (Nash, 1951)**.** *Every finite strategic game (where each player has finitely many actions) has at least one mixed strategy Nash equilibrium.*

A few important things follow from this:

Pure vs. Mixed Strategy NE

- Every pure strategy is a special mixed strategy (one action gets probability 1, all others get 0). So every pure strategy NE is also a mixed strategy NE.
- The reverse is not true: a mixed strategy NE where players genuinely randomize is not a pure strategy NE.
- Nash's theorem guarantees that even when no pure strategy NE exists (like Matching Pennies), a mixed strategy NE always does.

4.9 Examples

Opera/Ballet Game

Consider the following game between a Man (M) and a Woman (W):

	opera	**ballet**
opera	$1,4$	$0,0$
ballet	$0,0$	$4,1$

Let $p = \Pr(\text{W plays opera})$ and $q = \Pr(\text{M plays opera})$.
Pure strategy NE: $(opera, opera)$ and $(ballet, ballet)$.
Finding the MSNE:
For W to be indifferent between opera and ballet, we need M's mixing probability q to satisfy:

$$U_W(\text{opera}) = U_W(\text{ballet}) \implies 4q + 0(1-q) = 0 \cdot q + 1(1-q) \implies 4q = 1 - q \implies q = \frac{1}{5}.$$

For M to be indifferent between opera and ballet, we need W's mixing probability p to satisfy:

$$U_M(\text{opera}) = U_M(\text{ballet}) \implies 1 \cdot p + 0(1-p) = 0 \cdot p + 4(1-p) \implies p = 4 - 4p \implies p = \frac{4}{5}.$$

MSNE: W plays opera with probability $\frac{4}{5}$, M plays opera with probability $\frac{1}{5}$.
Expected payoffs in equilibrium:

$$U_W = 4q = 4 \cdot \frac{1}{5} = \frac{4}{5}, \qquad U_M = p = \frac{4}{5}.$$

So this game has three Nash equilibria in total: the two pure strategy NE and the mixed strategy NE $\left(\frac{4}{5}, \frac{1}{5}\right)$.

Game of Chicken

Luigi and Brad drive toward each other. Each can **Swerve** or **Not Swerve**. If one swerves and the other does not, the non-swerver wins \$10 and the swerver loses \$10. If both swerve, neither wins or loses. If neither swerves, they crash: Luigi loses \$80 and Brad loses \$120.

			Brad
		Swerve	Not Swerve
Luigi	Swerve	0 0	−10 10
	Not Swerve	10 −80	−10 −120

Pure strategy NE: (Swerve, Not Swerve) and (Not Swerve, Swerve).

Let p = Pr(Luigi plays Not Swerve) and q = Pr(Brad plays Not Swerve).

Find q so Luigi is indifferent:

$$U_L(\text{Swerve}) = U_L(\text{Not Swerve}) \implies 0 \cdot (1-q) + (-10)q = 10(1-q) + (-80)q$$

$$-10q = 10 - 10q - 80q = 10 - 90q \implies 80q = 10 \implies q = \frac{1}{8}.$$

Find p so Brad is indifferent:

$$U_B(\text{Swerve}) = U_B(\text{Not Swerve}) \implies 0(1-p) + (-10)p = -10(1-p) + (-120)p$$

$$-10p = -10 + 10p - 120p = -10 - 110p \implies 100p = 10 \implies p = \frac{1}{10}.$$

MSNE: Luigi plays Not Swerve with probability $\frac{1}{10}$, Brad plays Not Swerve with probability $\frac{1}{8}$.

Intuition: Brad has a nicer car so crashing is more costly for him. This makes him less aggressive in equilibrium ($q = \frac{1}{8} < \frac{1}{10}$). But because Brad is less likely to be aggressive, Luigi can afford to be slightly more aggressive. Each player's equilibrium mixing is determined entirely by the *other* player's payoffs.

Part 2: Brad's damage falls to $30.

The payoff matrix changes only in the (Not Swerve, Not Swerve) cell for Brad, from −120 to −30. Redoing Brad's indifference condition:

$$-10p = -10 - (30 - 10)p = -10 - 20p \implies 10p = 10 \implies p = \frac{1}{2}.$$

Luigi's indifference condition is unchanged since Luigi's payoffs did not change, so q still satisfies:

$$-10q = 10 - 90q \implies q = \frac{1}{8}.$$

Wait, Luigi's payoffs did not change so his indifference gives the same $q = \frac{1}{8}$. But now $p = \frac{1}{2}$: Luigi becomes far more aggressive. The intuition is that when Brad's car is cheaper, crashing is less catastrophic for Brad, so Brad is less deterred from not swerving. Luigi, knowing Brad is less scared of a crash, must randomize more aggressively to keep Brad indifferent.

4.10 Dominated Actions

In the context of mixed strategies, the notion of dominance extends naturally. A pure action can be dominated not just by another pure action, but by a **mixed strategy**.

Strict and Weak Dominance by Mixed Strategies

Player i's mixed strategy α_i **strictly dominates** action a_i' if:

$$U_i(\alpha_i, a_{-i}) > u_i(a_i', a_{-i}) \quad \text{for every } a_{-i}.$$

Player i's mixed strategy α_i **weakly dominates** action a_i' if the above holds with $\geq$ for all a_{-i} and $>$ for at least one.

Theorem 4.2. *A strictly dominated action is never played with positive probability in any mixed strategy Nash equilibrium.*

This is a useful result: before solving for a MSNE, you can first eliminate strictly dominated actions and work with the reduced game. Consider the following example:

Dominance by a Mixed Strategy

	X	Y	Z
A	$(5, 0)$	$(1, 7)$	$(2, -2)$
B	$(0, 4)$	$(4, 1)$	$(3, -1)$
C	$(-3, -3)$	$(-2, -4)$	$(-1, -5)$

For Player 1, action C is strictly dominated. Compare C to the mixed strategy that plays A and B each with probability $\frac{1}{2}$:

Against X: $\frac{1}{2}(5) + \frac{1}{2}(0) = 2.5 > -3$.

Against Y: $\frac{1}{2}(1) + \frac{1}{2}(4) = 2.5 > -2$.

Against Z: $\frac{1}{2}(2) + \frac{1}{2}(3) = 2.5 > -1$.
The mixed strategy $(\frac{1}{2}A, \frac{1}{2}B)$ strictly dominates C in every case. For Player 2, action Z yields a negative payoff in every column and is dominated as well. So neither C nor Z will appear in any MSNE.

4.11 Interpretations of Mixed Strategy Nash Equilibrium

A natural question: do players actually flip coins and randomize in real life? The standard interpretation has two criticisms. First, players in practice do not seem to consciously randomize. Second, even if they did, why would they randomize in a way that makes their opponent indifferent?

Harsanyi (1973) resolved this with his **Purification Theorem**. The idea is that the payoffs in a game are an approximation of reality. In practice, each player has small private inclinations that slightly nudge their payoffs one way or another, and these are not observed by others. Because of this private information, each player ends up choosing a unique best response almost all the time. But from an outside observer's perspective, not knowing these private shocks, the behavior looks like randomization.

Formally, if we add small independent payoff shocks $\varepsilon \cdot \eta_{ia}$ to each player's payoffs, the resulting game with incomplete information has a pure strategy equilibrium where each player follows a threshold rule. As $\varepsilon \to 0$, these pure strategy equilibria converge to the mixed strategy equilibria of the original game.

Harsanyi's Theorem

- To the **decision maker**: the choice is deterministic (she picks a strict best response given her private shock).
- To the **outside observer**: the behavior appears probabilistic, since the shocks are unobserved.

Mixed strategies are therefore not about players actually randomizing. They describe the uncertainty of an outside observer about which action a player will take.

Chapter 5: Dynamic Games, Backward Induction, and Subgame Perfection

5.1 Extensive Form

Normal form games model **simultaneous** decisions. But many strategic situations are **sequential** (one player moves, the other observes, then responds.)

An extensive form game consists of:

1. A set of **players** (including possibly **Nature** for random events).
2. A **tree**, a set of nodes and directed branches where each node has at most one incoming branch and any two nodes are connected by a unique path.
3. An assignment of each decision node to a player.
4. An **information partition**.
5. **Payoffs** at each terminal node.

Definition: Information Set

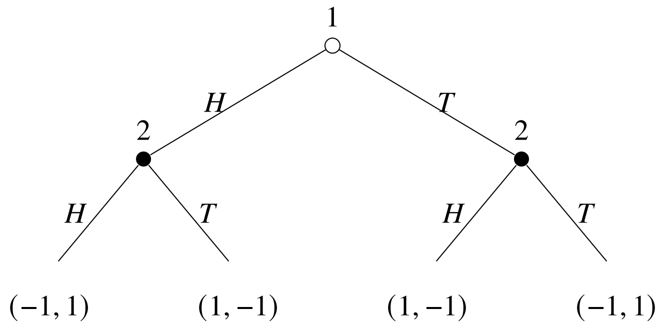
An **information set** is a collection of nodes $\{n_1, \ldots, n_k\}$ such that the same player moves at each node and the same actions are available at each node. The player knows they are somewhere in the set but **cannot distinguish** which node exactly.
A game where *all* information sets are singletons is called a **game of perfect information**, every player sees the full history before moving.

Definition: Strategy

A **strategy** for player i is a **complete contingent plan**, it specifies an action at *every* information set where player i might have to move, including those never reached on the equilibrium path.

Example: Matching Pennies with Perfect Information

Player 1 moves first, Player 2 *observes* Player 1's choice and then moves.



Player 2 **knows** what Player 1 chose, each of Player 2's nodes is its own singleton information set. Player 2's strategies are therefore complete contingent plans:

HH = Head regardless, HT = match Player 1, TH = opposite, TT = Tail regardless.

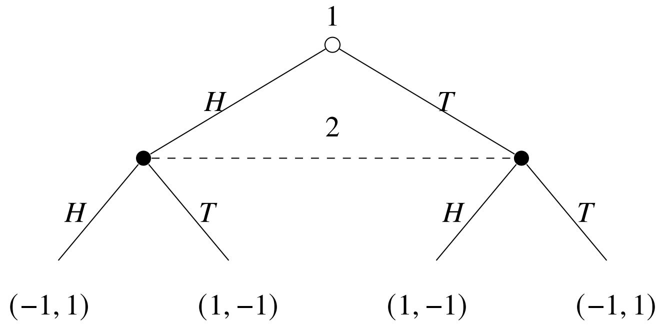
The normal form of this game is:

		Player 2			
		HH	HT	TH	TT
Player 1	H	$(-1, 1)$	$(-1, 1)$	$(1, -1)$	$(1, -1)$
	T	$(1, -1)$	$(-1, 1)$	$(1, -1)$	$(-1, 1)$

The Nash equilibria are (H, TH) and (T, HT) — in both cases Player 2 matches Player 1, and Player 1 is indifferent.

Example: Matching Pennies with Imperfect Information

Same setup, but now Player 2 does **not** observe Player 1's choice. Both of Player 2's decision nodes belong to the *same* information set (shown by the dashed oval).



Because both nodes are in the same information set, Player 2 cannot condition on Player 1's action. Player 2 now has only **two** strategies: H or T. The normal form collapses to the familiar 2×2 Matching Pennies matrix, this game is strategically identical to the simultaneous version.

Note: The same tree can represent very different games depending on the information structure. Perfect information → more strategies for the later mover; imperfect information → fewer strategies, and the game may reduce to the simultaneous case.

5.2 Backward Induction

The problem with Nash equilibrium in sequential games is that it allows **non-credible threats**. Consider the Entry Game: an Incumbent threatens to Fight if a Challenger enters, so the Challenger stays Out. But if the Challenger *did* enter, the Incumbent would prefer to Acquiesce (getting 1 rather than 0). The threat is empty, yet Nash equilibrium cannot rule it out.

Backward induction fixes this by requiring rational behavior at *every* node, even those off the equilibrium path.

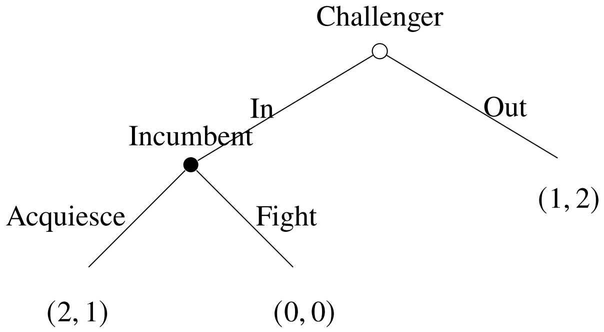
Backward Induction Procedure

Start at the nodes just before terminal nodes. The player there picks the action giving the highest payoff. Replace that subtree with the resulting payoff vector. Move up the tree and repeat until the root is reached.

The key assumption: it is **Note** that every player acts rationally at every node, even nodes that will not be reached under the equilibrium strategies.

Example: The Entry Game

An Incumbent is in a market. A Challenger decides to enter (In) or stay out (Out). If the Challenger enters, the Incumbent chooses to **Acquiesce** or **Fight**.



Payoffs are (Challenger, Incumbent). The normal form is:

		Incumbent	
		Acquiesce	Fight
Challenger	In	(2, 1)	(0, 0)
	Out	(1, 2)	(1, 2)

This game has **two Nash equilibria**: $(In, Acquiesce)$ and $(Out, Fight)$.
The equilibrium $(Out, Fight)$ survives as a Nash equilibrium because if the Challenger plays Out, the Incumbent never actually has to move, so threatening to Fight costs nothing. But this threat is **not credible**: if the Challenger entered, the Incumbent would strictly prefer Acquiesce $(1 > 0)$.

Applying backward induction:

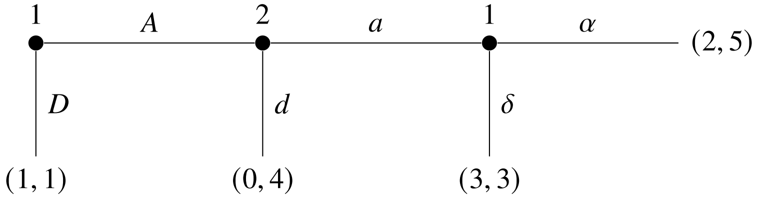
At the Incumbent's node: Acquiesce gives 1, Fight gives 0, so the Incumbent chooses **Acquiesce**. Replace the Incumbent's subtree with $(2, 1)$.

The Challenger now faces: In → 2, Out → 1, so the Challenger chooses **In**.

The unique backward induction outcome is $(In, Acquiesce)$ with payoffs $(2, 1)$. The non-credible threat is eliminated.

Example: The Centipede Game

Two players interact over three periods. At each node the active player can go **Across** (continue) or **Down** (exit). Payoffs grow the longer they stay in, but each player is tempted to exit just before the other does.



Applying backward induction:

Step 1 (Player 1's last node): α gives 2, δ gives 3. Player 1 chooses δ. Replace with $(3, 3)$.

Step 2 (Player 2's node): a leads to $(3, 3)$ so Player 2 gets 3; d gives 4. Player 2 chooses d. Replace with $(0, 4)$.

Step 3 (Player 1's first node): A leads to $(0, 4)$ so Player 1 gets 0; D gives 1. Player 1 chooses D.

The backward induction outcome is (D, d, δ) with payoffs $(1, 1)$.

This result is striking. Both players could get $(3, 3)$ or even $(2, 5)$ if they cooperated, but because neither can **commit** to continuing, the game collapses on the very first move. This illustrates the power of commitment, and the cost of its absence.

Example: Stackelberg Duopoly

Two firms compete in quantities. Unlike Cournot, here **Firm 1 (leader) moves first** and **Firm 2 (follower) observes** q_1 before choosing q_2.

Inverse demand: $P(Q) = \alpha - Q$ where $Q = q_1 + q_2$, marginal cost $c > 0$, $\alpha > c$.

Step 1: Solve Firm 2's problem (backward induction starts here).

Firm 2 maximizes $\pi_2 = (\alpha - q_1 - q_2 - c)\, q_2$ taking q_1 as given:

$$\frac{\partial \pi_2}{\partial q_2} = \alpha - q_1 - 2q_2 - c = 0 \implies BR_2(q_1) = \frac{\alpha - c - q_1}{2}$$

Step 2: Solve Firm 1's problem, anticipating Firm 2's reaction.

Firm 1 substitutes $BR_2(q_1)$ into its own profit and maximizes:

$$\pi_1 = \left(\alpha - q_1 - \frac{\alpha - c - q_1}{2} - c\right) q_1 = \frac{(\alpha - c - q_1)}{2}\, q_1$$

$$\frac{\partial \pi_1}{\partial q_1} = \frac{\alpha - c - 2q_1}{2} = 0 \implies q_1^* = \frac{\alpha - c}{2}$$

Step 3: Find Firm 2's equilibrium quantity.

$$q_2^* = BR_2(q_1^*) = \frac{\alpha - c - \frac{\alpha - c}{2}}{2} = \frac{\alpha - c}{4}$$

Comparing Stackelberg and Cournot outcomes:

	Cournot	**Stackelberg**
Firm 1	$\frac{\alpha - c}{3}$	$\frac{\alpha - c}{2}$
Firm 2	$\frac{\alpha - c}{3}$	$\frac{\alpha - c}{4}$
Total Q	$\frac{2(\alpha - c)}{3}$	$\frac{3(\alpha - c)}{4}$

The leader produces *more* and the follower produces *less* than in Cournot. By moving first and committing to a high quantity, the leader forces the follower to scale back, this is the **first-mover advantage**.

Note: there exists another Nash equilibrium where Firm 2 threatens to produce the Cournot quantity regardless of Firm 1's choice. But this is a **non-credible threat**, once Firm 1 has chosen q_1^*, Firm 2's best response is strictly q_2^*, not the Cournot quantity. Backward induction rules this out.

5.3 Subgame Perfection

Backward induction works well for games of perfect information. But what about sequential games where some information sets contain more than one node? We cannot apply backward induction directly there. The right generalization is **Subgame Perfect Equilibrium (SPE)**.

Definition: Subgame

A **subgame** consists of a single decision node n, all nodes that follow n, and all branches connecting them, such that: if an information set has one node inside the subgame, it has *all* its nodes inside the subgame.

Every game is a subgame of itself. Subgames other than the full game are called **proper subgames**.

Definition: Subgame Perfect Equilibrium

A Nash equilibrium is a **Subgame Perfect Equilibrium** if and only if it induces a Nash equilibrium in *every* subgame of the game.

Note: SPE rules out equilibria that rely on non-credible threats or promises at any point in the game, even off the equilibrium path. Every SPE is a Nash equilibrium, but not every Nash equilibrium is subgame perfect:

$$\{\text{Subgame Perfect Equilibria}\} \subseteq \{\text{Nash Equilibria}\}$$

To find SPE we use the same logic as backward induction: solve the innermost proper subgames first, replace them with their equilibrium payoffs, and work backwards toward the root.

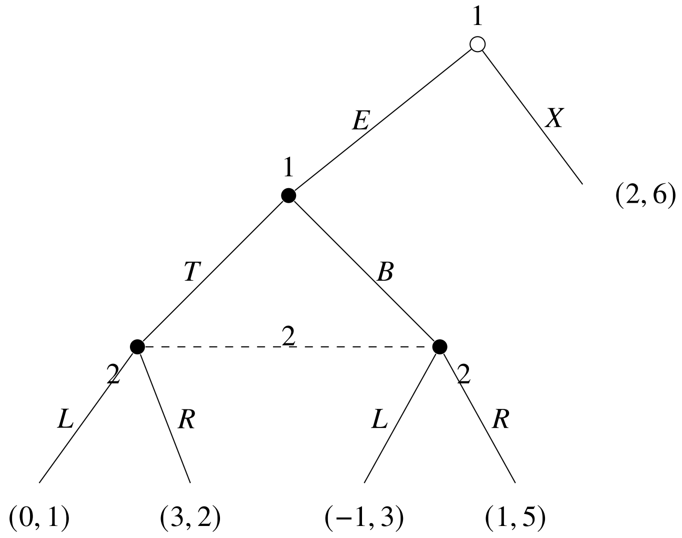
Proposition

For finite perfect information games, the set of subgame perfect equilibria is exactly the set of strategy profiles found by backward induction.

Every finite extensive form game with perfect recall has at least one subgame perfect equilibrium.

Example: SPE with Imperfect Information

Player 1 moves first choosing E (enter the subgame) or X (exit). If Player 1 chooses E, Player 1 then moves again choosing T or B. Player 2 moves after that but **does not observe** whether Player 1 chose T or B (imperfect information — dashed oval). Player 2 chooses L or R.



How many subgames? There are **two**: the proper subgame starting after Player 1 plays E (the part involving Player 1's choice of T/B and Player 2's response), and the whole game itself. Note that Player 2's information set spans both nodes after T and B — the subgame must contain both nodes together, which it does.

Step 1: Solve the proper subgame (the part after E).

This subgame is itself a simultaneous game between Player 1 (choosing T or B) and Player 2 (choosing L or R, without observing Player 1's choice). The payoff matrix of this subgame is:

		Player 2	
		L	R
Player 1	T	$(0, 1)$	$(3, 2)$
	B	$(-1, 3)$	$(1, 5)$

Finding the Nash equilibrium of this subgame:

Player 1 checks best responses. If Player 2 plays L: Player 1 prefers T ($0 > -1$). If Player 2 plays R: Player 1 prefers T ($3 > 1$). So T **strictly dominates** B for Player 1.

Given Player 1 plays T, Player 2 prefers R ($2 > 1$).

The unique Nash equilibrium of this subgame is (T, R) with payoffs $(3, 2)$. **Step 2: Replace the proper subgame with its equilibrium payoff** $(3, 2)$ **and solve the remaining game.**

Player 1 now faces: $E \longrightarrow 3$ or $X \longrightarrow 2$. Player 1 chooses **E**.

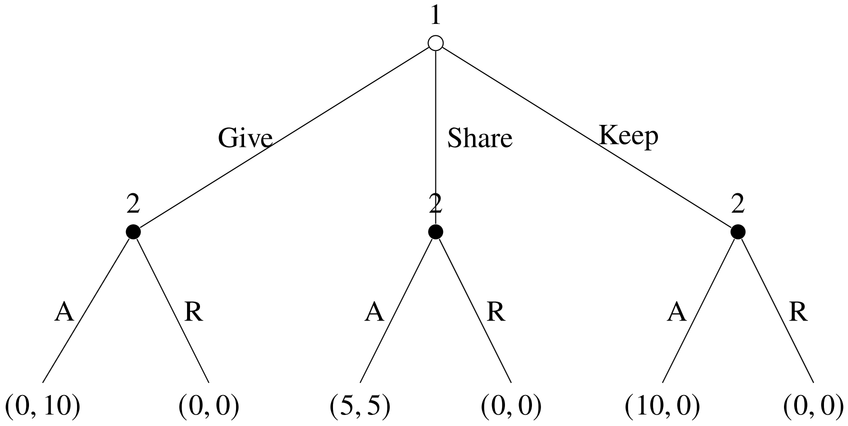
The unique SPE is: Player 1 plays E then T; Player 2 plays R. Equilibrium payoffs: $(3, 2)$. **Is there a Nash equilibrium that is NOT subgame perfect here?**

Yes. Consider the strategy profile: Player 1 plays X, Player 1 would play T inside the subgame, Player 2 plays L. Given Player 2 plays L, the subgame payoff to Player 1 is 0, making X ($\longrightarrow 2$) optimal. Given Player 1 plays X, Player 2's strategy does not affect the payoff so L is vacuously optimal. This *is* a Nash equilibrium. But it is **not** subgame perfect: inside the proper subgame,

L is strictly dominated for Player 2 (given T dominates B for Player 1, Player 2 gets 2 from R and 1 from L). Player 2 would never actually play L if the subgame were reached. The threat is non-credible.

Example: The Ultimatum Game

Player 1 has \$10 and proposes a split. She can **Give** (\$0 for herself, \$10 for Player 2), **Share** (\$5 each), or **Keep** (\$10 for herself, \$0 for Player 2). Player 2 observes the offer and can **Accept** or **Reject**. If accepted, payoffs are as proposed. If rejected, both get \$0.



How many subgames? There are **4**: one proper subgame starting at each of Player 2's three nodes (after Give, Share, Keep), plus the whole game.

Step 1: Solve each of Player 2's subgames.

- After **Give**: Accept → \$10, Reject → \$0. Player 2 **Accepts**.
- After **Share**: Accept → \$5, Reject → \$0. Player 2 **Accepts**.
- After **Keep**: Accept → \$0, Reject → \$0. Player 2 is **indifferent** — both Accept and Reject are best responses.

Step 2: Solve Player 1's problem.

Case 1 — Player 2 Accepts after Keep:

Player 1 gets: Give → \$0, Share → \$5, Keep → \$10. Player 1 chooses **Keep**.

SPE: Player 1 plays Keep; Player 2 Accepts everything. Payoffs: $(10, 0)$.

Case 2 — Player 2 Rejects after Keep:

Player 1 gets: Give → \$0, Share → \$5, Keep → \$0. Player 1 chooses **Share**.

SPE: Player 1 plays Share; Player 2 Accepts Give and Share, Rejects Keep. Payoffs: $(5, 5)$.

Both are valid SPE. The threat to reject Keep is credible in Case 2 because Player 2 is indifferent between Accept and Reject after Keep (\$0 either way) — it costs Player 2 nothing to follow through on this threat.

What is NOT a SPE? Any equilibrium where Player 2 rejects Share is not subgame perfect, since after a Share offer, Reject gives \$0 while Accept gives \$5. That rejection threat is strictly non-credible.

5.4 Existence and Backward Induction

Proposition: The set of subgame perfect equilibria of a finite perfect information game is exactly the set of strategy profiles found by backward induction.

Every finite extensive form game with perfect recall has at least one subgame perfect equilibrium.

5.5 Examples

Synergistic Relationships

Two individuals choose effort levels **sequentially**: Player 1 chooses $a_1 \geq 0$ first, then Player 2 observes a_1 and chooses $a_2 \geq 0$. Player i's payoff is:

$$u_i(a_i, a_j) = a_i(c + a_j - a_i), \quad c > 0$$

Step 1: Solve Player 2's problem.
Player 2 maximizes $u_2 = a_2(c + a_1 - a_2)$ taking a_1 as given:

$$\frac{\partial u_2}{\partial a_2} = c + a_1 - 2a_2 = 0 \implies BR_2(a_1) = \frac{c + a_1}{2}$$

Step 2: Solve Player 1's problem, anticipating Player 2's reaction.
Player 1 substitutes $BR_2(a_1)$ into her own payoff:

$$u_1 = a_1\left(c + \frac{c + a_1}{2} - a_1\right) = a_1 \cdot \frac{3c + a_1 - 2a_1}{2} = a_1 \cdot \frac{3c - a_1}{2}$$

Maximizing over a_1:

$$\frac{\partial u_1}{\partial a_1} = \frac{3c - 2a_1}{2} = 0 \implies a_1^* = \frac{3c}{2}$$

Step 3: Find Player 2's equilibrium effort.

$$a_2^* = BR_2(a_1^*) = \frac{c + \frac{3c}{2}}{2} = \frac{5c}{4}$$

Equilibrium payoffs:

$$u_1^* = \frac{3c}{2}\left(c + \frac{5c}{4} - \frac{3c}{2}\right) = \frac{3c}{2} \cdot \frac{3c}{4} = \frac{9c^2}{8}$$

$$u_2^* = \frac{5c}{4}\left(c + \frac{3c}{2} - \frac{5c}{4}\right) = \frac{5c}{4} \cdot \frac{5c}{4} = \frac{25c^2}{16}$$

Note that Player 2 benefits from moving second, by observing Player 1's effort and best responding, Player 2 free-rides on the synergy Player 1 creates. This is the **second-mover advantage** in synergistic settings, in contrast to the first-mover advantage we saw in the Stackelberg model.

Chapter 6: Repeated Games

6.1 Introduction

Think about the interactions you have with people you see repeatedly, a neighbor, a coworker, a business partner. In those situations, you might cooperate even when you could get away with not doing so, because you know you will face them again. This is repeated games: when players interact over and over, the **shadow of the future** can discipline behavior that would otherwise break down.

6.2 Preferences: The Discount Factor

When a game is repeated over time, each player receives a stream of payoffs (one per period.) To evaluate that stream, we need to know how the player weighs the future relative to today.

Each player i has a **discount factor** $\delta \in (0, 1)$, and evaluates the sequence of outcomes $(a^1, a^2, \dots, a^T)$ by the **discounted sum**:

$$u_i(a^1) + \delta\, u_i(a^2) + \delta^2\, u_i(a^3) + \cdots + \delta^{T-1} u_i(a^T) = \sum_{t=1}^{T} \delta^{t-1} u_i(a^t)$$

The discount factor δ captures patience. A player with δ close to 0 barely cares about the future, she is **very impatient**. A player with δ close to 1 values the future almost as much as the present, she is **very patient**. We assume all players share the same discount factor δ throughout.

Intuition: a payoff of x received one period from now is worth only δx today. Two periods from now, it is worth $\delta^2 x$. The further away a payoff is, the less it matters right now.

Example: Discounting a Payoff Stream

Suppose a player earns 3 every period forever. Her discounted sum is:

$$3 + 3\delta + 3\delta^2 + \cdots = \frac{3}{1-\delta}$$

This uses the geometric series formula: $1 + \delta + \delta^2 + \cdots = \frac{1}{1-\delta}$ for $\delta \in (0, 1)$. If instead she deviates and earns 5 today then 1 every period after:

$$5 + \delta + \delta^2 + \cdots = 5 + \frac{\delta}{1-\delta}$$

Cooperation is better when $\frac{3}{1-\delta} \geq 5 + \frac{\delta}{1-\delta}$, which simplifies to $\delta \geq \frac{1}{2}$.

6.3 Equivalent Payoff Functions

You already know that preferences over deterministic outcomes can be represented by any strictly increasing transformation of a utility function, the ranking is what matters, not the numbers themselves. For lotteries, the same holds for increasing affine transformations of the Bernoulli utility function. What happens with sequences of payoffs?

The answer turns out to be the same: only affine transformations preserve preferences over streams.

Lemma: Equivalent Payoff Functions

Suppose there are at least three possible outcomes. The discounted sum of payoffs with payoff function u and discount factor δ represents the **same preferences over payoff streams** as the discounted sum with payoff function v and the same δ if and only if there exist constants α and $\beta > 0$ such that

$$u(x) = \alpha + \beta\, v(x) \quad \text{for all } x.$$

In other words, you can shift and scale the payoff function freely, but you cannot apply a nonlinear transformation and expect the preferences over streams to be preserved. This is more restrictive than the single-outcome case, and it matters whenever we want to compare payoff functions across different representations of the same game.

The Discounted Average

If preferences over streams are represented by the discounted sum $\sum_{t=1}^{\infty} \delta^{t-1} u(a^t)$, they are equally well represented by the **discounted average**:

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u(a^t)$$

This is just a positive scaling by $(1-\delta)$, so it represents the same preferences. But it has a nice interpretive advantage: its values are directly comparable to single-period payoffs.

To see why, notice that if a player earns the same payoff x in every period, her discounted average is exactly x:

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} x = (1-\delta) \cdot \frac{x}{1-\delta} = x$$

6.4 Infinitely Repeated Games

In an **infinitely repeated game**, the players interact in the same strategic game over and over, with no fixed final period. There is always a tomorrow.

If the end date is very far in the future, or if players are uncertain about exactly when the game will end, treating the horizon as infinite is often a good approximation of how players actually reason. They focus on the ongoing relationship, not on counting down to the last period.

Formally, let G be a strategic game with player set N, where each player i has action set A_i and payoff function u_i. The **infinitely repeated game** of G with discount factor δ is an extensive game with perfect information and simultaneous moves in which:

- The set of players is N.
- The set of terminal histories is the set of all infinite sequences $(a^1, a^2, \ldots)$ of action profiles from G.
- After every non-terminal history, every player i chooses an action from A_i — the same action set as in the stage game G.
- Each player i evaluates each terminal history $(a^1, a^2, \ldots)$ by the discounted average:

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

The game G that is repeated each period is called the **stage game**. At the start of each period, both players simultaneously choose actions, observe the outcome, and move on to the next period. Because information is perfect, the full history of past play is always observed before each new period begins.

A **history** up to period t is simply the sequence of action profiles played so far: $(a^1, a^2, \ldots, a^t)$. The empty history $\emptyset$ refers to the very start of the game, before anyone has moved.

6.5 Strategies in a Repeated Game

A strategy in a repeated game is the same idea as in any extensive game: a **complete contingent plan** specifying what to do after every possible history, including histories that would never be reached if both players follow their strategies. Formally, a strategy for player i specifies an action $a_i \in A_i$ for every finite sequence $(a^1, \ldots, a^T)$ of past outcomes of G, including the empty history $\emptyset$. In practice, since there are infinitely many possible histories, we describe strategies compactly by specifying the *rule* a player follows as a function of what has happened so far.

Three Important Strategies:

1. Grim Trigger

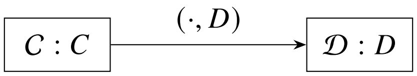
In grim trigger strategy, player i starts by cooperating and continues to do so as long as the other player has always cooperated. The moment the other player defects even once, player i defects *forever*, with no possibility of return to cooperation.

Formally, player i's grim trigger strategy s_i is defined as:

$$s_i(\emptyset) = C$$

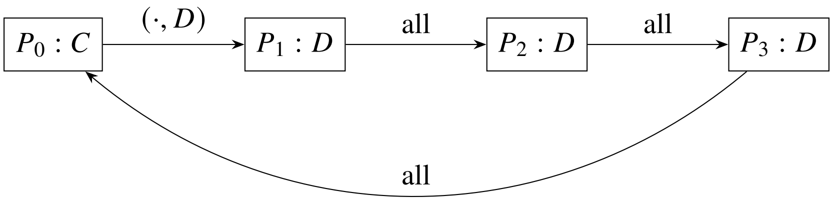
$$s_i(a^1, \ldots, a^t) = \begin{cases} C & \text{if } a_j^\tau = C \text{ for all } \tau = 1, \ldots, t \\ D & \text{otherwise} \end{cases}$$

We can picture this strategy as having two **states**. Initially the strategy is in state $\mathcal{C}$, where C is played. If the other player ever plays D, the strategy moves permanently to state $\mathcal{D}$, where D is played forever. Once in state $\mathcal{D}$, there is no arrow out.



2. Finite-Period Punishment

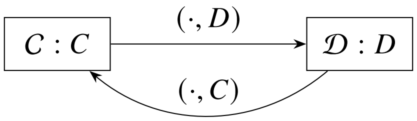
A less extreme strategy punishes deviations for only a fixed number of periods k, then returns to cooperation regardless of what happened during the punishment phase. The diagram below shows the case $k = 3$:



Notice that during the punishment phases P_1, P_2, P_3, the transition says “all outcomes” — meaning the punishment runs for exactly three periods no matter what the deviating player does during that time. After the third punishment period, cooperation resumes automatically.

3. Tit-for-Tat

The Tit-for-tat says *do whatever the other player did in the previous period.* Start by cooperating, then mirror the opponent’s last action exactly.



Unlike grim trigger, tit-for-tat *forgives*: if the deviating player returns to C, tit-for-tat immediately returns to C as well. The punishment length is endogenous, it lasts exactly as long as the other player keeps defecting. If the other player defects forever, so does tit-for-tat. If the other player returns to cooperation after one period, tit-for-tat does too.

Comparing the Three Strategies

Strategy	**Punishment length**	**Forgiveness?**
Grim Trigger	Infinite	Never
Finite Punishment (k periods)	Fixed k periods	Yes, after k periods
Tit-for-Tat	Depends on opponent	Yes, immediately

6.6 Infinitely Repeated Prisoner's Dilemma

Going back to the Prisoner's Dilemma stage game:

		Player 2	
		C	D
Player 1	C	2 2	0 3
	D	3 0	1 1

In the one-shot version, (D, D) is the unique Nash equilibrium. D strictly dominates C for each player, so cooperation collapses. But what about the infinitely repeated version? If players are sufficiently patient, can (C, C) be sustained every period?

The answer is yes, but only if δ is large enough, and only under the right strategies.

Grim Trigger as a Nash Equilibrium

Suppose both players adopt the grim trigger strategy: start with C, and defect forever the moment the other player defects. If both follow this strategy, the outcome is (C, C) every period, giving each player the stream $(2, 2, 2, \ldots)$, whose discounted average is simply 2.

Now suppose player 2 considers deviating. The best deviation is to play D as early as possible, say, in period 1. After that, player 1 will defect forever (grim trigger fires), and since D is the best response to D, player 2 will also defect forever. So the best possible deviation yields the stream $(3, 1, 1, 1, \ldots)$.

The discounted average of the deviation stream is:

$$(1-\delta)\left[3+\delta+\delta^2+\cdots\right] = (1-\delta)\left[3+\frac{\delta}{1-\delta}\right] = 3(1-\delta)+\delta$$

Player 2 cannot gain by deviating if and only if the cooperation payoff is at least as large as the deviation payoff:

$$2 \geq 3(1-\delta)+\delta$$

$$2 \geq 3-3\delta+\delta = 3-2\delta$$

$$2\delta \geq 1 \quad \Rightarrow \quad \delta \geq \frac{1}{2}$$

By symmetry, the same condition applies to player 1. We conclude:

Grim Trigger Nash Equilibrium

If $\delta \geq \frac{1}{2}$, then the strategy pair in which both players use the grim trigger strategy is a Nash equilibrium of the infinitely repeated Prisoner's Dilemma. Cooperation (C, C) occurs in every period.

Tit-for-Tat as a Nash Equilibrium

Now consider the strategy pair where both players use tit-for-tat. As before, if player 2 can gain by deviating at all, she can gain by deviating in period 1, playing D. After that, player 1 switches to D in period 2, then mirrors whatever player 2 does. So player 2 faces a choice between two best deviations:

Option 1: Alternate between D and C. After deviating with D in period 1, player 1 plays D in period 2. If player 2 reverts to C in period 2, player 1 returns to C in period 3, and the situation resets. This generates the alternating stream $(3, 0, 3, 0, \ldots)$. Its discounted average is:

$$(1 - \delta) \cdot \frac{3}{1 - \delta^2} = \frac{3}{1 + \delta}$$

Option 2: Defect every period. This gives the stream $(3, 1, 1, \ldots)$ with discounted average $3(1 - \delta) + \delta = 3 - 2\delta$.

Since the cooperation payoff under tit-for-tat is 2, tit-for-tat is a best response to tit-for-tat if and only if both deviations are unprofitable:

$$2 \geq \frac{3}{1 + \delta} \quad \text{and} \quad 2 \geq 3 - 2\delta$$

Both conditions simplify to $\delta \geq \frac{1}{2}$. So:

Tit-for-Tat Nash Equilibrium

If $\delta \geq \frac{1}{2}$, then the strategy pair in which both players use tit-for-tat is also a Nash equilibrium of the infinitely repeated Prisoner's Dilemma.

Examples

The following three examples use a different Prisoner's Dilemma payoff matrix, where C means *confess* and NC means *not confess*. They explore what happens when the punishment lasts for different numbers of periods.

	C	**NC**
C	$2,2$	$6,0$
NC	$0,6$	$3,3$

Here the cooperative outcome is (NC, NC) with payoff 3 each, the stage-game Nash equilibrium is (C, C) with payoff 2 each, and deviating from cooperation gives a one-period payoff of 6.

Example 1: Punishment for One Period

Question. Can players support (NC, NC) as a SPNE using tit-for-tat strategies that punish a deviation by reverting to (C, C) for just *one* period, then returning to cooperation?

Cooperation payoff stream:

$$3 + 3\delta + 3\delta^2 + \cdots$$

Deviation payoff stream (deviate today, punished for one period, then back to cooperation):

$$\underbrace{6}_{\text{gain}} + \underbrace{2\delta}_{\text{punishment}} + \underbrace{3\delta^2 + 3\delta^3 + \cdots}_{\text{back to cooperation}}$$

Cooperation can be sustained if the cooperation stream is at least as large as the deviation stream:

$$3 + 3\delta + 3\delta^2 + \cdots \geq 6 + 2\delta + 3\delta^2 + \cdots$$

The $3\delta^2 + \cdots$ terms cancel from both sides, leaving:

$$3 + 3\delta \geq 6 + 2\delta \quad \Rightarrow \quad \delta \geq 3$$

Since δ must satisfy $0 < \delta < 1$, this condition can never hold. **Cooperation cannot be sustained with only one period of punishment.**

Intuition: the gain from deviating today $(6 - 3 = 3)$ is much larger than the loss from one period of punishment $(3 - 2 = 1)$. A single period of punishment is simply not enough of a deterrent.

Example 2: Punishment for Two Periods

Question. Can (NC, NC) be supported if players revert to (C, C) for *two* periods after a deviation, then return to cooperation?

Cooperation payoff stream:

$$3 + 3\delta + 3\delta^2 + 3\delta^3 + \cdots$$

Deviation payoff stream:

$$\underbrace{6}_{\text{gain}} + \underbrace{2\delta + 2\delta^2}_{\text{punishment}} + \underbrace{3\delta^3 + \cdots}_{\text{back to cooperation}}$$

Cooperation requires:

$$3 + 3\delta + 3\delta^2 + 3\delta^3 + \cdots \geq 6 + 2\delta + 2\delta^2 + 3\delta^3 + \cdots$$

The $3\delta^3 + \cdots$ terms cancel, leaving:

$$3 + 3\delta + 3\delta^2 \geq 6 + 2\delta + 2\delta^2$$

$$\delta + \delta^2 \geq 3 \quad \Rightarrow \quad \delta^2 + \delta - 3 \geq 0$$

Solving using the quadratic formula:

$$\delta = \frac{-1 \pm \sqrt{1 + 12}}{2} = \frac{-1 \pm \sqrt{13}}{2}$$

This gives $\delta_1 = -2.3 < 0$ and $\delta_2 = 1.3 > 1$. Since neither root lies in $(0, 1)$, the condition $\delta^2 + \delta - 3 \geq 0$ cannot be satisfied for any valid discount factor. **Cooperation still cannot be sustained with two periods of punishment.**

Two periods of punishment is better than one, but the gain from deviating today still outweighs the discounted future losses.

Example 3: Punishment for Ten Periods

Question. Suppose players revert to (C, C) for *ten* periods after a deviation. Find the threshold discount factor δ above which cooperation can be sustained.

Cooperation payoff stream:

$$3 + 3\delta + 3\delta^2 + \cdots + 3\delta^{10} + 3\delta^{11} + \cdots$$

Deviation payoff stream:

$$6 + 2\delta + 2\delta^2 + \cdots + 2\delta^{10} + 3\delta^{11} + \cdots$$

Cooperation requires:

$$3 + 3\delta + \cdots + 3\delta^{10} + 3\delta^{11} + \cdots \geq 6 + 2\delta + \cdots + 2\delta^{10} + 3\delta^{11} + \cdots$$

The $3\delta^{11} + \cdots$ terms cancel. Rearranging the remaining terms:

$$(3 - 2)\delta + (3 - 2)\delta^2 + \cdots + (3 - 2)\delta^{10} \geq 6 - 3$$

$$\delta + \delta^2 + \cdots + \delta^{10} \geq 3$$

$$\delta \left(1 + \delta + \delta^2 + \cdots + \delta^9\right) \geq 3$$

After discarding solutions outside $(0, 1)$, the numerical solution gives:

$$\delta \geq 0.76$$

Cooperation can be sustained with ten periods of punishment, provided $\delta \geq 0.76$. Comparing the three examples: one period requires $\delta \geq 3$ (impossible), two periods requires $\delta \geq 1.3$ (impossible), and ten periods requires $\delta \geq 0.76$ (feasible). The longer the punishment, the lower the patience threshold needed — but players still need to care enough about the future for any finite punishment to work.

6.7 Examples

The general approach in every repeated game problem follows three steps:

1. Compute the discounted payoff stream from **cooperating** forever.
2. Find the **optimal deviation** and compute its payoff stream (deviation gain today, then punishment forever after).
3. Find the condition on δ under which cooperation dominates deviation.

Collusion in a Cournot Duopoly

Two firms compete as Cournot oligopolists in a market with inverse demand:

$$p(q_1, q_2) = a - bq_1 - bq_2$$

Both firms have total cost $TC(q_i) = cq_i$ where $c > 0$ and $a > c$. The firms interact repeatedly with discount factor δ.

Step 1: Cournot (Stage-Game Nash Equilibrium).
Each firm maximizes its own profit taking the other's output as given. Firm 1 solves:

$$\max_{q_1} \ (a - bq_1 - bq_2)q_1 - cq_1$$

Taking the first-order condition:

$$\frac{\partial \pi_1}{\partial q_1} = a - 2bq_1 - bq_2 - c = 0 \quad \Rightarrow \quad B_1(q_2) = \frac{a-c}{2b} - \frac{1}{2}q_2$$

By symmetry, $B_2(q_1) = \frac{a-c}{2b} - \frac{1}{2}q_1$. Solving the system by setting $q_1 = q_2 = q^*$:

$$q^* = \frac{a-c}{2b} - \frac{1}{2}q^* \quad \Rightarrow \quad \frac{3}{2}q^* = \frac{a-c}{2b} \quad \Rightarrow \quad q_1^* = q_2^* = \frac{a-c}{3b}$$

The market price and each firm's Cournot profit are:

$$p^{\text{cournot}} = \frac{a+2c}{3}, \qquad \pi_i^{\text{cournot}} = \frac{(a-c)^2}{9b}$$

Step 2: Cartel (Cooperative) Outcome.
If both firms collude, they jointly maximize total profit as if they were a single monopolist. Let

$Q = q_1 + q_2$:

$$\max_{Q} \ (a - bQ)Q - cQ$$

First-order condition: $a - 2bQ - c = 0 \ \Rightarrow \ Q^{\text{cartel}} = \frac{a-c}{2b}$, so each firm produces:

$$q_1^{\text{cartel}} = q_2^{\text{cartel}} = \frac{a-c}{4b}$$

The cartel price and each firm's profit are:

$$p^{\text{cartel}} = \frac{a+c}{2}, \qquad \pi_i^{\text{cartel}} = \frac{(a-c)^2}{8b}$$

Since $\frac{(a-c)^2}{8b} > \frac{(a-c)^2}{9b}$, both firms earn strictly higher profits in the cartel than under Cournot competition. This is why they have an incentive to collude — but also why each has an incentive to secretly deviate.

Step 3: Optimal Deviation.

Suppose firm 2 is producing the cartel output $q_2^{\text{cartel}} = \frac{a-c}{4b}$ and firm 1 considers deviating. Firm 1 plugs firm 2's output into its best response function:

$$q_1^{\text{dev}} = \frac{a-c}{2b} - \frac{1}{2} \cdot \frac{a-c}{4b} = \frac{a-c}{2b} - \frac{a-c}{8b} = \frac{3(a-c)}{8b}$$

Firm 1's deviation profit (with firm 2 still at cartel output) is:

$$\pi_1^{\text{dev}} = \left[a - b \cdot \frac{3(a-c)}{8b} - b \cdot \frac{a-c}{4b} - c\right] \frac{3(a-c)}{8b} = 3\left(\frac{a-c}{8}\right)\left(\frac{3(a-c)}{8b}\right) = \frac{9(a-c)^2}{64b}$$

Since $\frac{9(a-c)^2}{64b} > \frac{(a-c)^2}{8b}$, firm 1 has a short-run incentive to deviate — it earns more today by producing more than the cartel agreement allows.

Step 4: Incentives to Cooperate.

Suppose the punishment for deviation is reversion to Cournot competition forever (grim trigger). The discounted payoff streams are:

Cooperate forever:

$$\frac{1}{1-\delta} \cdot \frac{(a-c)^2}{8b}$$

Deviate today, then Cournot forever:

$$\frac{9(a-c)^2}{64b} + \frac{\delta}{1-\delta} \cdot \frac{(a-c)^2}{9b}$$

Cooperation is sustained when:

$$\frac{1}{1-\delta} \cdot \frac{(a-c)^2}{8b} \geq \frac{9(a-c)^2}{64b} + \frac{\delta}{1-\delta} \cdot \frac{(a-c)^2}{9b}$$

Dividing through by $\frac{(a-c)^2}{b}$ and simplifying:

$$\frac{1}{8(1-\delta)} \geq \frac{9}{64} + \frac{\delta}{9(1-\delta)}$$

Multiplying through by $72(1-\delta)$:

$$9 \geq \frac{9 \cdot 72(1-\delta)}{64} + 8\delta \quad \Rightarrow \quad \delta \geq \frac{9}{17}$$

Result

The cartel agreement can be sustained as a Nash equilibrium of the infinitely repeated Cournot game if and only if

$$\delta \geq \frac{9}{17} \approx 0.529.$$

Firms need to be sufficiently patient — they must value future cartel profits enough that the short-run gain from deviating is not worth the long-run cost of triggering Cournot competition.

Are the punishments credible?

If firm 2 keeps producing the cartel output while firm 1 deviates, firm 2 earns $\frac{3(a-c)^2}{32b}$. But if firm 2 switches to its Cournot output, it earns $\frac{(a-c)^2}{9b}$. Since $\frac{(a-c)^2}{9b} > \frac{3(a-c)^2}{32b}$ for all parameter values, firm 2 strictly prefers to revert to Cournot upon observing a deviation. The punishment is credible.

6.8 One-Shot Deviation and Grim Trigger

Nash equilibrium does not require behavior to be rational at every history, only along the equilibrium path. For subgame perfection we need rationality everywhere, but in an infinitely repeated game there are infinitely many subgames to check. The **One-Shot Deviation Principle** gives us a shortcut.

Theorem: One-Shot Deviation Principle

A strategy profile in an infinitely repeated game is a subgame perfect equilibrium if and only if no player can gain by deviating from her strategy in **just one period**, after any history, while following her strategy in all subsequent periods.

Grim Trigger is a SPNE: Suppose both players use grim trigger. Two types of history to check.

Type 1: Histories where grim trigger prescribes C.

Following the strategy gives discounted average 2. Deviating once to D gives stream $(3, 1, 1, \ldots)$ with discounted average:

$$(1-\delta)\left[3+\delta+\delta^2+\cdots\right] = 3(1-\delta)+\delta$$

No profitable deviation if and only if:

$$2 \geq 3(1-\delta)+\delta \quad \Rightarrow \quad \delta \geq \frac{1}{2}$$

Type 2: Histories where grim trigger prescribes D.

Following the strategy gives discounted average 1. Deviating once to C gives stream $(0, 1, 1, \ldots)$ with discounted average:

$$(1-\delta)\left[0+\delta+\delta^2+\cdots\right] = \delta$$

No profitable deviation if and only if $1 \geq \delta$, which always holds. Both cases are satisfied when $\delta \geq \frac{1}{2}$:

Result: Grim Trigger is a SPNE

If $\delta \geq \dfrac{1}{2}$, then grim trigger is a **subgame perfect Nash equilibrium** of the infinitely repeated Prisoner's Dilemma. The punishment is credible because in the punishment phase (D, D) is itself the stage-game Nash equilibrium, playing D is a best response to D, so there is no temptation to forgive.