Module 4 — Information And Beliefs

Game Theory

hbar.university

Bayesian Games

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.1

A Bayesian game (Harsanyi, 1967–68) is a game of **incomplete information**: players have private information about their preferences, their capabilities, or the state of the world. Harsanyi's insight was that any such situation can be modeled as a game in which nature draws a **type** for each player at the start, each player privately observes their own type, and then players act. This "Harsanyi transformation" converts incomplete information into imperfect information within a larger game tree, which can then be analyzed using the standard tools of Module 1–3. The node introduces the formal apparatus — types, type spaces, common priors — and sets up Bayesian Nash equilibrium (4.3) as the solution concept.

Formal Definition

A **Bayesian game** is a tuple

$$G^B = (N, \{\Theta_i\}, \{S_i\}, p, \{u_i\}),$$

where:

- N is the finite set of players.
- Θ_i is the set of possible **types** for player i. A type $\theta_i \in \Theta_i$ is a complete specification of player i's private information (preferences, beliefs, capabilities). The type profile is $\theta = (\theta_1, \dots, \theta_n) \in \Theta = \prod_i \Theta_i$.
- S_i is the set of actions available to player i (assumed independent of type for notational simplicity; can be generalized).
- $p \in \Delta(\Theta)$ is the **common prior** distribution over type profiles. It is common knowledge among players.
- $u_i : S \times \Theta \to \mathbb{R}$ is player i's payoff, which depends on the full action profile and the full type profile.

Type-contingent strategy. A pure strategy for player i is a function $s_i : \Theta_i \to S_i$: a plan of what action to take for each possible realization of own type. Mixed strategies are probability distributions over such functions, or equivalently type-contingent behavioral strategies.

Interim belief. Given own type θ_i, player i's belief about others' types is the conditional distribution $p(\theta_{-i} \mid \theta_i)$, derived from the common prior via Bayes' rule:

$$p(\theta_{-i} \mid \theta_i) = \frac{p(\theta_i, \theta_{-i})}{\sum_{\theta'_{-i}} p(\theta_i, \theta'_{-i})}.$$

Interim expected payoff. Given own type θ_i and a type-contingent strategy profile s, player i's expected payoff is

$$U_i(s \mid \theta_i) = \sum_{\theta_{-i}} p(\theta_{-i} \mid \theta_i) \cdot u_i(s_i(\theta_i), s_{-i}(\theta_{-i}), \theta_i, \theta_{-i}).$$

The **Harsanyi transformation** represents this as a game of imperfect information: a chance node at the root draws $\theta \sim p$, each player observes only θ_i, then players choose actions.

Intuition

Incomplete information is the rule, not the exception. In any realistic strategic situation, players know things about themselves that others don't — their valuations in an auction, their costs in a market, their private signals about the state of the world.

Harsanyi's move was to *absorb this incompleteness into the game specification.* Instead of saying "player 1 doesn't know player 2's type," we say "there is a chance move at the start that draws types, and player 1 observes only their own type." This reduces incomplete information to imperfect information — a tree with chance nodes and information sets — and the standard tools apply.

The common prior is a strong assumption: it says all players agree on the distribution of types before observing their own type. They may update this to different posterior beliefs after observing their types, but the starting point is common knowledge. Under the common prior, differences in beliefs arise only from differences in information, not from differences in priors.

A type can encode any private information: preferences (how much you value a good), capabilities (your cost function), or beliefs about the state of the world (what you think nature has done). Harsanyi showed that all of these reduce to a type space, provided you're willing to make the type space big enough.

Structural Role

Bayesian games are the ontological foundation for every subsequent incomplete-information topic:

- **4.1 ⇒ 4.2.** Types and type spaces are formalized further; hierarchies of beliefs.
- **4.1 ⇒ 4.3.** Bayesian Nash equilibrium is the solution concept.
- **4.1 ⇒ 4.4.** Perfect Bayesian equilibrium extends Bayesian Nash to sequential games.
- **4.1 ⇒ 4.5.** Signaling games are two-stage Bayesian games.
- **4.1 ⇒ 6.1.** Mechanism design is the inverse problem: design the game so Bayesian Nash equilibria yield desired outcomes.
- **4.1 ⇒ 6.3.** Auction theory uses Bayesian games to model private valuations.
- **4.1 ⇒ 4.6.** Sequential equilibrium is the belief-consistent refinement.

The introduction of types is also the formal entry point for poker game theory: each player's hole cards are their type, the common prior is the uniform deal, and the resulting Bayesian game is the full poker game.

Mathematical Development

The Harsanyi transformation

Start with incomplete information: player i has payoff function $u_i(\cdot, \theta_i)$ that depends on a private parameter $\theta_i \in \Theta_i$. Player j does not know θ_i. Instead of "player j has no idea," assume player j has a *probabilistic belief* $p_j(\theta_i)$.

Under a common prior, all players share the same joint distribution $p(\theta)$, with $p_j(\theta_i)$ the appropriate marginal. The Bayesian game is then:

1. Chance draws $\theta \sim p$ at the root.
2. Player i observes θ_i (and nothing else).
3. Players choose actions simultaneously (or sequentially, for extensive-form Bayesian games).
4. Payoffs are $u_i(s, \theta)$ at the leaves.

This is a game of imperfect information (the information sets group nodes differing in unseen θ_{-i}). The extensive-form machinery of Module 3 applies.

Type independence and correlated types

Independent types. $p(\theta) = \prod_i p_i(\theta_i)$. Each player's type is drawn independently. This is the simplest case and is standard in auction theory.

Correlated types. $p(\theta)$ does not factor. Correlation represents shared information, e.g., common-value auctions where all bidders have noisy signals about the same underlying value.

Conditional independence. Types are independent given some "state" variable, but marginal types are correlated.

Common priors and no-speculative-trade

Aumann's agreement theorem (1976). Under a common prior and common knowledge of the posterior, two Bayesian agents cannot "agree to disagree" about the expected value of any random variable. More strongly, they must have the same posterior.

No-speculative-trade theorem (Milgrom–Stokey, 1982). In a market with common priors, no trades can occur solely on the basis of private information — because every trade requires a buyer and seller whose beliefs diverge from the common prior, which Aumann's theorem rules out.

These results highlight how strong the common-prior assumption is. Dropping it gives a different theory (with "belief heterogeneity").

The type space as belief hierarchy

A rich type for player i includes not just their own payoff parameter but also their belief about others' parameters, their belief about others' beliefs about their parameter, and so on — an infinite hierarchy. Harsanyi showed that this hierarchy can always be encoded as a point in some finite-dimensional type space (the "universal type space"), provided beliefs are coherent. Mertens–Zamir (1985) formalized this and proved existence of a universal type space.

For most applications, one uses a finite type space with types chosen to capture the relevant private information directly (e.g., Θ_i = set of hole-card combinations in poker).

Worked Example

A simple Bayesian coordination game

Two players choose $\{A, B\}$. Player 1's payoff from coordination is known to be 2; Player 2's payoff from coordination depends on Player 2's type. Player 2 has two types, $\theta_2 \in \{H, L\}$, with $p(H) = 0.4, p(L) = 0.6$.

Player 2's payoff: - Type H: $u_2(A, A) = 3, u_2(B, B) = 1, u_2(\text{mismatch}) = 0$. - Type L: $u_2(A, A) = 1, u_2(B, B) = 3, u_2(\text{mismatch}) = 0$.

Player 1's payoff: $u_1(A, A) = 2, u_1(B, B) = 2, u_1(\text{mismatch}) = 0$. (Independent of Player 2's type.)

Player 1's interim belief. $p(H \mid \theta_1) = 0.4$ since θ_1 is empty (no private info) — the prior itself.

Player 1's expected payoff from A (supposing Player 2 plays best response given their type): Player 2 type H plays A (prefers matching A = 3 to mismatching = 0); type L plays B (prefers matching B = 3). So Player 1's expected payoff from A is $0.4 \cdot 2 + 0.6 \cdot 0 = 0.8$. From B: $0.4 \cdot 0 + 0.6 \cdot 2 = 1.2$. Player 1 plays B.

Bayesian Nash equilibrium: Player 1 plays B, Player 2 plays $s_2(H) = A, s_2(L) = B$. Verify: given Player 1 plays B, Player 2 type H gets 0 from A and 1 from B — type H should switch to B. Re-iterate.

Corrected analysis: given Player 1 plays B, type H best response is B (payoff $1 > 0$), type L best response is B (payoff $3 > 0$). So Player 2's strategy is $s_2 = (B, B)$. Player 1's best response to both types playing B: expected payoff from B is 2, from A is 0 — play B. Confirmed equilibrium: $(B, (B, B))$, payoffs $(2, \text{type-weighted})$.

Auction as a Bayesian game

First-price sealed-bid auction with two bidders. Each bidder has a private valuation $v_i \sim U[0, 1]$, drawn independently. Bidders submit bids simultaneously; higher bid wins and pays their bid.

- Types: $\Theta_i = [0, 1]$, $p(v) = \text{Uniform}[0, 1]^2$.
- Strategy: $s_i : [0, 1] \to [0, \infty)$, the bid as a function of valuation.
- Payoff: $u_i(b_1, b_2, v_i) = (v_i - b_i) \cdot \mathbb{1}[b_i > b_{-i}]$.

Bayesian Nash equilibrium (derived in Node 6.3): symmetric bidding with $s_i(v_i) = v_i/2$ for two bidders (and $v_i \cdot (n - 1)/n$ for n bidders). Each bidder shades their bid below their valuation because winning with a higher bid yields lower surplus.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Type θ_i	Your hole cards — the private information you hold
Type space Θ_i	$\binom{52}{2} = 1326$ possible 2-card combinations
Common prior $p(\theta)$	Uniform over all valid joint dealings of hole cards (no two players holding the same card)
Interim belief $p(\theta_{-i} \mid \theta_i)$	Your "range-the-opponent" reasoning: given you hold AK, what are the possible hands they have?
Type-contingent strategy $s_i(\theta_i)$	Your range strategy: a plan specifying what to do with each possible hole-card combination
Common prior assumption	Everyone at the table agrees the deck is uniformly shuffled

The "type" is literally your hole cards. When poker players talk about "playing your range" or "thinking in ranges," they are thinking about the Bayesian game structure: each player has a private type drawn from a type space, plays a type-contingent strategy, and the opponents do not know the realization. Harsanyi's transformation is what makes this a well-defined game theory problem.

Concrete situation

Flop decision as a Bayesian decision problem. You are BTN, facing a flop check from BB on $Kd9s4c$.

- **Your type:** your specific hole cards, say $A\clubsuit K\clubsuit$.
- **Opponent's type space:** all hole card combinations consistent with BB's preflop calling range.
- **Common prior on opponent's type (given their preflop action):** the distribution over combinations in BB's call-preflop range.

- **Interim belief:** your posterior over BB's hole cards, updated for the fact that they checked the flop (which for typical solver ranges includes a fairly wide check-call range).
- **Your strategy:** a type-contingent rule — "with AKcc, bet 33%; with 72o, give up / check" — a specific function from your hole-card type to an action.

Bayesian Nash structure. At equilibrium, your strategy is a best response to BB's strategy given your interim beliefs, and vice versa. The joint fixed point is the Bayesian Nash equilibrium of the poker game (restricted to the subgame rooted at this flop spot with these input ranges). The solver computes this fixed point.

What this understanding buys you

"Thinking in ranges" is Bayesian game-theoretic thinking. When a coach tells you "don't think about their hand, think about their range," they are translating the Bayesian game framework into poker jargon. Your opponent's type is private; the strategic object of interest is the *distribution* over their types, and you respond to the distribution, not to a specific realization.

Bayesian updating is the core move. Every opponent action updates your posterior over their type. Preflop, you have a prior (random 2 cards or filtered by position). They raise — update. They call your 3-bet — update. They check-call the flop — update. Each action narrows their range by Bayesian conditioning, which is exactly $p(\theta_{-i} \mid \text{history})$ computed via Bayes' rule.

You are also a type from their perspective. The symmetry is crucial: you are thinking about their range, but they are simultaneously thinking about yours. In equilibrium, each of you is best-responding to the other's type-contingent strategy, and the joint profile is a fixed point. Both ranges are common knowledge in the common-prior sense; what's hidden is the type realizations.

Common prior is a real assumption. If you and opponent disagree about "what the preflop range is," you are not in a common-prior Bayesian game. This is a form of belief heterogeneity and can create exploitable gaps. Most solver study is based on a shared assumption about what ranges "should" look like in equilibrium — violating this (by playing non-standard ranges) is a form of exploitation.

Cross-Links

- **1.1 Players, strategies, payoffs** — the base ontology.
- **3.2 Information sets** — the extensive-form representation of types.
- **4.2 Types and type spaces** — formalization of private information.
- **4.3 Bayesian Nash equilibrium** — the solution concept.
- **4.4 Perfect Bayesian equilibrium** — sequential refinement.
- **4.5 Signaling games** — two-stage Bayesian games.
- **6.1, 6.3 Mechanism design, auctions** — applied Bayesian games.
- **Statistics — Bayesian inference** — the mathematical toolkit.

Structural Compression

Invariant: A Bayesian game is a normal-form game enriched with a type space for each player and a common prior over type profiles, representing incomplete information as a lottery over strategic situations; the Harsanyi transformation converts any such situation into a standard imperfect-information extensive-form game.

Mechanism: Type chance node at the root, private type observation by each player, type-contingent strategies, and interim Bayesian belief computations over opponents' types.

Cross-System Echo: Bayesian games are the game-theoretic instantiation of the general pattern "uncertain parameters observed privately by agents": hidden-variable models in statistics, partial-observation settings in reinforcement learning, and private-state channels in information theory all instantiate the same information asymmetry with an appropriate prior.

Exercises

1. Write down the Bayesian game structure for a first-price sealed-bid auction with two bidders and private uniform valuations.
2. Compute the interim belief of player i of type θ_i about θ_{-i} in a Bayesian game with a given common prior.
3. Consider a coordination game where Player 2 has two equally likely types, each with different preferred actions. Find the Bayesian Nash equilibrium.
4. Show that under a common prior, two Bayesian agents with the same information must have the same posterior (Aumann agreement).
5. Construct a Bayesian game with correlated types and write the conditional distribution $p(\theta_{-i} \mid \theta_i)$.
6. Formulate 2-card heads-up hold'em with uniform deal as a Bayesian game. Specify Θ_i, p, and the payoff function at showdown.
7. **Argue, structurally,** why the Harsanyi transformation is a "free" reduction: every game of incomplete information can be represented as a game of imperfect information with a chance move, without loss of strategic content. Identify what additional assumption — beyond the type space itself — is required to make the transformation work.

Types and Type Spaces

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.2

Node 4.1 introduced the Bayesian game as a tuple with a set of types Θ_i for each player. This node zooms in on the type object itself: what a type encodes, how it relates to hierarchies of beliefs about beliefs, and when two different-looking type spaces are strategically equivalent. The key insight, due to Harsanyi and formalized by Mertens–Zamir (1985), is that any coherent hierarchy of beliefs collapses into a point in a **universal type space**. For applied work — including poker — one almost always uses a finite type space that directly encodes payoff-relevant private information. This node makes precise what "a type" contains and sets up Bayesian Nash equilibrium (4.3) as equilibrium over type-contingent strategies.

Formal Definition

Let N be the set of players, and for each player i let Θ_i be a measurable set of **types**. The product $\Theta = \prod_{i \in N} \Theta_i$ is the set of **type profiles**.

Payoff-parameter type. The simplest type encodes player i's private payoff parameter: a type θ_i indexes a payoff function $u_i(\cdot, \theta_i)$. Opponents do not observe θ_i.

Belief-hierarchy type. A richer type also encodes player i's beliefs about others. Define the hierarchy recursively:

- **Level 0:** X_i^0 = the payoff-relevant state for player i (their payoff parameter, private signal).
- **Level 1:** $X_i^1 = X_i^0 \times \Delta(X_{-i}^0)$ — own parameter plus a belief about others' parameters.
- **Level** $k+1$**:** $X_i^{k+1} = X_i^k \times \Delta(X_{-i}^k)$ — own parameter, belief about others' levels up to k.

A **belief hierarchy** for player i is an element $(\mu_i^1, \mu_i^2, \ldots)$ with $\mu_i^k \in \Delta(X_{-i}^{k-1})$, subject to **coherence**: each μ_i^{k+1} has marginal on X_{-i}^{k-1} equal to μ_i^k.

Type space. A type space is a tuple $(\Theta_i, \hat{\theta}_i, \pi_i)_{i \in N}$ where each type $\theta_i \in \Theta_i$ is associated with:

- $\hat{\theta}_i(\theta_i) \in X_i^0$: the payoff parameter implied by type θ_i.
- $\pi_i(\theta_i) \in \Delta(\Theta_{-i})$: an interim belief about the other players' types.

Iterating these two maps generates a belief hierarchy from θ_i. The type space is **universal** if every coherent belief hierarchy is represented by exactly one type.

Theorem (Mertens–Zamir, 1985). The universal type space exists and is unique up to isomorphism. Every type space embeds into it.

Common prior. A type space has the **common prior property** if there exists a single distribution $p \in \Delta(\Theta)$ such that each $\pi_i(\theta_i)$ equals the conditional $p(\theta_{-i} \mid \theta_i)$ derived from p.

Intuition

Two distinct questions live inside "what player 2 doesn't know about player 1":

1. **What is player 1's payoff parameter?** (their valuation, their hand, their cost.)
2. **What does player 1 believe about player 2's payoff parameter?** And what does player 1 believe player 2 believes player 1 believes? And so on.

Question 1 is simple: player 1 has a parameter, player 2 has a distribution over it. Question 2 is the Russian-doll one: beliefs about beliefs about beliefs, stretching infinitely. Harsanyi's genius was to notice that *you can fold the whole tower into a single "type" object* — a point in a sufficiently rich space — so that knowing the type tells you both the payoff parameter and the beliefs at every level.

Why does this work? Because the beliefs at level $k+1$ are distributions over level-k beliefs, each belief at every level is eventually a distribution over the original payoff parameters. Coherence (the marginals must match) forces the hierarchy to be consistent. Under coherence, the hierarchy is fully determined by a single probability measure, and that measure is what the "type" represents.

The common prior assumption is a separate and strong restriction. Without it, each player can have an arbitrary interim belief $\pi_i(\theta_i)$; with it, all interim beliefs are consistent with a single joint distribution that players would have agreed on before observing their types. Aumann's agreement theorem (1976) and no-speculative-trade rely on this property.

For applied work, you almost never write down the belief hierarchy explicitly. Instead, you pick a finite type space where each θ_i is literally a payoff parameter (a hole card combination, a valuation, a cost), and the hierarchy is generated automatically from the common prior. This is fine: the Mertens–Zamir theorem guarantees that what you lose is only redundant representation, not strategic content.

Structural Role

Types and type spaces are the bridge between Bayesian games as tuples (4.1) and Bayesian Nash equilibrium (4.3):

- **4.1 ⇒ 4.2.** Bayesian games are defined; this node unpacks the type object.
- **4.2 ⇒ 4.3.** Bayesian Nash equilibrium is equilibrium over type-contingent strategies.
- **4.2 ⇒ 4.4.** Perfect Bayesian equilibrium extends to sequential games with belief updating.
- **4.2 ⇒ 4.5.** Signaling games use types where the sender's type is private and the receiver updates beliefs from signals.
- **4.2 ⇒ 4.6.** Sequential equilibrium imposes consistency on beliefs at off-path information sets.
- **4.2 ⇒ 6.1, 6.3.** Mechanism design and auction theory rest on a type space for private valuations.

The universal type space result is mostly of theoretical importance — it guarantees the foundational framework is well-posed. Applied game theory, including every poker solver, uses small finite type spaces.

Mathematical Development

Coherence of belief hierarchies

A belief hierarchy $(\mu_i^1, \mu_i^2, \ldots)$ is **coherent** if for every k,

$$\text{marg}_{X_{-i}^{k-1}} \mu_i^{k+1} = \mu_i^k.$$

Incoherent hierarchies are rejected: they would correspond to a player believing contradictory things at different levels. Under coherence, the sequence of beliefs converges to a single probability measure on a suitable limit space, by Kolmogorov's extension theorem.

Brandenburger–Dekel (1993). The set of coherent hierarchies, with an additional "common knowledge of coherence" requirement, is homeomorphic to a compact space — the universal type space. Every type space maps into it continuously.

Finite type spaces vs universal type space

A **finite type space** uses Θ_i = finite set (e.g., $\{H, L\}$ for "high-cost" vs "low-cost" player), with a specified common prior $p \in \Delta(\Theta)$. The induced hierarchy is derived by iterating π_i.

The universal type space is infinite-dimensional and useful for existence theorems and foundational work. In applied settings, one argues “pick a sufficiently rich finite type space for the payoff-relevant parameters, assume common prior, and compute.” This gives all the strategic content of the universal space for the problem at hand.

When finite type spaces suffice. If the payoff parameter takes finitely many values and players know the distribution, a finite type space with types indexing parameter values and common prior equal to the parameter distribution is strategically equivalent to the full universal type space (for this problem).

Common prior property and its strength

The common prior assumption is logically separate from the definition of types. A type space without common prior has types that specify interim beliefs directly; there need not be any joint prior consistent with all of them.

Why common prior is strong. Aumann argued (1987) that rational agents with different priors are behaving as if they had different information, and there is no way to distinguish prior from posterior without referring to some earlier state. He therefore advocated common prior as a modeling default. Critics argue that prior heterogeneity is real and economically important (Morris 1995; asset pricing with disagreement).

In game theory practice, most models assume common prior. Departures are flagged as “belief heterogeneity” or “non-common-prior types.”

Type space morphisms and equivalence

Two type spaces are **strategically equivalent** if they induce the same set of Bayesian Nash equilibrium outcome distributions for any game played on them. Equivalence classes of type spaces embed uniquely into the universal type space. This is the sense in which “the universal type space is universal”: it represents every equivalence class exactly once.

Different authors use different finite type spaces for the same problem (e.g., different poker solver ranges); as long as the payoff-parameter structure and common prior agree, the equilibria are the same.

Interim, ex ante, and ex post

Three vantage points on the Bayesian game:

- **Ex ante (before types are drawn).** Expected payoffs averaged over own type and opponents’ types, under the common prior.
- **Interim (after own type is observed but before play).** Expected payoffs averaged over opponents’ types only, conditional on own type.
- **Ex post (after play is realized).** The actual realized payoffs.

Bayesian Nash equilibrium is typically defined at the interim stage: each type of each player must be playing a best response given the interim belief. This is equivalent to ex ante optimality when the common prior is used.

Worked Example

A two-player Bayesian game with finite types

Player 1 has type $\theta_1 \in \{H, L\}$; Player 2 has type $\theta_2 \in \{H, L\}$. Common prior:

$\theta_1 \setminus \theta_2$	H	L
H	0.3	0.1
L	0.1	0.5

Marginal distributions. $p(\theta_1 = H) = 0.4, p(\theta_1 = L) = 0.6$; similarly for θ_2.

Interim beliefs.

$$p(\theta_2 = H \mid \theta_1 = H) = \frac{0.3}{0.4} = 0.75, \quad p(\theta_2 = L \mid \theta_1 = H) = 0.25.$$

$$p(\theta_2 = H \mid \theta_1 = L) = \frac{0.1}{0.6} = 0.167, \quad p(\theta_2 = L \mid \theta_1 = L) = 0.833.$$

The types are **positively correlated**: knowing you are H raises your probability that the opponent is H from 0.4 to 0.75.

Belief hierarchy (level 2). Given $\theta_1 = H$, Player 1's belief about Player 2's belief about θ_1 is derived by asking: *given Player 2's type is H (prob 0.75 from Player 1's view), what does Player 2 believe about θ_1?* Player 2 of type H thinks $p(\theta_1 = H \mid \theta_2 = H) = 0.3/0.4 = 0.75$, and similarly Player 2 of type L thinks $p(\theta_1 = H \mid \theta_2 = L) = 0.1/0.6 = 0.167$. So Player 1's level-2 belief is a mixture $0.75 \cdot$ (Player 2 of type H's belief) + $0.25 \cdot$ (Player 2 of type L's belief), all derived automatically from the common prior via iterated Bayes' rule.

The entire infinite hierarchy is generated from the 4-entry common prior table. This is the compactness of types: infinite hierarchy, finite representation.

Non-common-prior types

Suppose instead Player 1's interim belief $\pi_1(H \mid \theta_1 = H) = 0.9$ and Player 2's interim belief $\pi_2(H \mid \theta_2 = H) = 0.5$. Can these come from a common prior?

Attempted common prior: let $p(H,H) = a$, $p(H,L) = b$, $p(L,H) = c$, $p(L,L) = d$. Then $0.9 = a/(a+b) \Rightarrow a = 9b$, and $0.5 = a/(a+c) \Rightarrow a = c$. So $a = 9b = c$, and normalization $a+b+c+d = 1 \Rightarrow 9b+b+9b+d = 1 \Rightarrow 19b + d = 1$. This is one equation in two unknowns — there exist such (b,d) satisfying it (e.g., $b = 0.05, d = 0.05, a = 0.45, c = 0.45$).

So these interim beliefs are compatible with a common prior. This illustrates that common prior is not automatic from interim beliefs, but is a testable constraint — and when satisfied, the common prior may still be non-unique.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Type θ_i	Your specific hole cards — the unique realization of your private information
Type space Θ_i	The set of hole-card combinations in your range — e.g., all hands BTN opens preflop
Payoff parameter	How strong your hand is against the opponent's range — your equity
Common prior	The deck is uniform, and both players agree on it (standard modeling assumption)
Interim belief $\pi_i(\theta_{-i} \mid \theta_i)$	Your posterior over the opponent's hole cards given everything observed so far, including the cards you hold
Belief hierarchy	"I think they have AK, and I think they think I have a bluff, and I think they think I think they're bluffing, . . . "

Formal object	Poker instantiation
Finite vs universal type space	Solvers use finite type spaces (bucketed ranges); universal type space is theoretical

Concrete situation

Range interpretation as type space. When you load a solver and see a preflop range chart like "BTN opens 22+, A2s+, A5o+, K9s+, KTo+, Q9s+, QTo+, J9s+, JTo, T9s, 98s, 87s, 76s, 65s," you are specifying a type space for the BTN player. Each element of the range is a possible type. The common prior is the uniform distribution over the deal, and the BTN's "type" is drawn from this range by Bayes' conditioning on the observation that BTN opened (since some hands fold and some open).

Card removal and correlation. Suppose BTN opens and BB calls with range R_{BB}. You (BTN) hold $A♠K♠$. What is your interim belief about BB's hand?

Two effects: 1. **Range filtering.** BB's hand is in R_{BB}. 2. **Card removal.** BB's hand cannot contain $A♠$ or $K♠$.

The interim belief is $p(\theta_{BB} \mid \theta_{BTN} = A♠K♠)$ = uniform over R_{BB} minus combinations containing the blockers, renormalized. This is the formal object the solver uses internally — Bayes' rule on hole cards conditional on observed actions and own cards.

Belief hierarchy in action. Consider "level reasoning" in poker: - Level 0: "I have AK, I'm strong." - Level 1: "My opponent has range R, I should play best response to R." - Level 2: "My opponent is thinking about my range R', they will play best response to R'." - Level 3: "My opponent thinks I think they have R, so they will deviate from R by bluffing more." - …

Each level is a belief about beliefs. In equilibrium (Bayesian Nash), all levels are consistent: your strategy is best-response to their strategy given your belief, and their strategy is best-response to yours given their belief, and these beliefs are the common-prior posteriors. The universal type space is the formal setting in which this consistency is well-defined.

Finite solver type space = universal in practice. For a specific spot (specific board, specific preflop history), the solver uses a finite type space per player (the preflop calling range, which has ~100–500 combos). This is strategically equivalent to the universal type space for the relevant payoff-relevant information. The Mertens–Zamir theorem guarantees nothing is lost.

What this understanding buys you

Type = hole cards is exact, not metaphor. When you read "in a Bayesian game, player i's type encodes their private information," and you're a poker player, the correct interpretation is literal: your type is $A♠K♠$, period. Every other object in the Bayesian game (type space, common prior, interim belief, type-contingent strategy) has a precise poker counterpart. The mathematics is the game.

Common prior is what "agreed ranges" means. GTO play assumes both players agree on what each player's preflop range is. This is a common prior on the joint type space. If you and your opponent disagree about what "BTN opens" or "BB defends" means, you are playing a non-common-prior game, and exploitation becomes possible (or impossible, depending on direction). This is the formal content of "the solver assumes equilibrium ranges."

Level reasoning collapses to equilibrium fixed point. "Level 1, level 2, level 3 thinking" is informal; the formal object is the Bayesian Nash equilibrium, which is the fixed point of all levels. Your interim strategy and opponent's interim strategy are jointly best-responses given the common prior. Solving for level-k thinking for finite k gives bounded rationality models (cognitive hierarchy, level-k models); as $k \to \infty$ and the beliefs become consistent, you converge to the Bayesian Nash equilibrium. This is what every solver computes.

Card removal is Bayesian conditioning, not a side effect. Poker players often treat "card removal" as a separate consideration ("don't forget blockers!"), but formally it is just the computation of $\pi_i(\theta_{-i} \mid \theta_i)$ under the common prior of a uniform deck. The blockers are automatically in the conditional distribution. When a solver reports your optimal frequencies, card-removal effects are already absorbed.

Cross-Links

- **4.1 Bayesian games** — the containing framework.
- **4.3 Bayesian Nash equilibrium** — the solution concept defined on type-contingent strategies.
- **4.4 Perfect Bayesian equilibrium** — dynamic version with belief updating on information sets.
- **4.5 Signaling games** — canonical two-stage Bayesian game with sender/receiver types.
- **4.6 Sequential equilibrium** — consistency of off-path beliefs.
- **6.1 Mechanism design** — design a game for a given type space.
- **6.3 Auction theory** — private valuations as types drawn from a prior.

Structural Compression

Invariant: A type encodes all payoff-relevant private information plus all levels of belief about other players' information; under coherence, an infinite hierarchy of beliefs collapses to a point in a universal type space, and any finite type space with a common prior is strategically equivalent to its embedding in that universal space.

Mechanism: Iterated belief-about-belief construction modulo coherence constraints; Bayesian conditioning on types at each level; common prior as a consistency condition tying all interim beliefs to a single joint distribution.

Cross-System Echo: Type spaces generalize the "state" variable in stochastic processes and POMDPs: a type is a sufficient statistic for a player's private information and all higher-order beliefs, just as a POMDP belief state is a sufficient statistic for the agent's knowledge about the hidden environment. The universal type space result parallels the Kolmogorov extension theorem for stochastic processes.

Exercises

1. For a two-player Bayesian game with $|\Theta_1| = |\Theta_2| = 2$, enumerate all coherent belief hierarchies at levels 1 and 2 and count them.
2. Show that under common prior, the level-k belief of player i can be derived from the common prior by k iterations of Bayes' rule.
3. Construct a two-player example where players have coherent interim beliefs but no common prior is compatible with them.
4. For the 2×2 common prior in the worked example, compute the level-2 belief of Player 1 of type H about Player 2's belief about θ_1.
5. Define strategic equivalence of type spaces formally and verify it for a simple example with two different finite type spaces encoding the same payoff structure.
6. In a poker context, compute the interim belief of BTN holding $A♠K♠$ about BB's hand, given BB's preflop calling range is $\{AK, AQ, AJ, KQ, KJ, QJ\}$ (all suits). Account for card removal.
7. **Argue, structurally,** why the universal type space result is necessary for the foundations of Bayesian game theory, even though applied work always uses finite type spaces. What would go wrong foundationally without it?

Bayesian Nash Equilibrium

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.3

Bayesian Nash equilibrium (BNE) is the solution concept for Bayesian games: a strategy profile in which every type of every player is playing a best response given their interim belief. It extends Nash equilibrium (2.1) to incomplete-information settings by making strategies type-contingent and best-response conditions interim. Harsanyi (1967–68) introduced BNE alongside the Bayesian game framework; it underwrites auction theory, mechanism design, and every incomplete-information application in game theory. For poker, Bayesian Nash equilibrium is what the solver computes — type = hole cards, strategy = what you do with each hand, and the equilibrium is a fixed point where both players are best-responding to each other's type-contingent strategies given the common prior.

Formal Definition

Let $G^B = (N, \{\Theta_i\}, \{S_i\}, p, \{u_i\})$ be a Bayesian game with common prior p.

Type-contingent strategy. A pure strategy is $s_i : \Theta_i \to S_i$. A behavioral (mixed) type-contingent strategy is $\sigma_i : \Theta_i \to \Delta(S_i)$.

Interim expected payoff. Given opponent strategy profile σ_{-i} and own type θ_i, player i's interim expected payoff from action $a \in S_i$ is

$$U_i(a \mid \theta_i, \sigma_{-i}) = \sum_{\theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{a_{-i}} \sigma_{-i}(\theta_{-i})(a_{-i}) \cdot u_i(a, a_{-i}, \theta_i, \theta_{-i}).$$

Bayesian Nash equilibrium. A strategy profile $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ is a **Bayesian Nash equilibrium** if, for every player i and every type $\theta_i \in \Theta_i$ with $p(\theta_i) > 0$,

$$\sigma_i^*(\theta_i) \in \arg\max_{\sigma_i \in \Delta(S_i)} U_i(\sigma_i \mid \theta_i, \sigma_{-i}^*).$$

Equivalently (by linearity): for every type θ_i and every action a in the support of $\sigma_i^*(\theta_i)$, a is a best response to σ_{-i}^* at θ_i.

Ex ante characterization. A profile is BNE iff each player's ex ante expected payoff (averaged over all types) is maximized by their strategy, holding opponents' strategies fixed. Under common prior, ex ante and interim optimality are equivalent at types in the support of p.

Theorem (existence). Every finite Bayesian game has a Bayesian Nash equilibrium. (Proof: treat the Bayesian game as a Nash equilibrium problem in a suitably defined "agent-normal-form" game where each type of each player is a separate pseudo-agent; apply Nash's existence theorem to the finite game on strategy profiles $\prod_{i,\theta_i} \Delta(S_i)$.)

Intuition

Think of each type of each player as a "separate pseudo-agent" with its own best-response problem. The common prior ties these agents together through the interim belief computation: when type θ_i of player i decides what to do, it computes an expected payoff by averaging over the opponent's types weighted by $p(\theta_{-i} \mid \theta_i)$.

A Bayesian Nash equilibrium is a profile where *every one* of these pseudo-agents is best-responding to the pseudo-agents of the other players. Equivalently: every type of every player is playing optimally given their beliefs, and no type wants to deviate.

Why “pseudo-agent”? Because the players are the real agents, but the action plan must account for every contingency (every type). Given the common prior, the whole profile can be computed all at once as a joint fixed point.

The notion of equilibrium here is *interim* — checked type by type. This is important because a player of type θ_i is not averaging over her own types (she knows her type); she is averaging only over opponents’ types. The ex ante formulation (averaging over everyone’s types) is equivalent under common prior but is a different computation — it ignores the conditional structure.

One conceptual subtlety: **BNE is just Nash equilibrium of the agent-normal-form game.** The Bayesian game, after the Harsanyi transformation, is a game of imperfect information with a chance node at the start. Its Nash equilibrium is the BNE. So BNE is not a new solution concept — it is Nash equilibrium applied to the right object.

Structural Role

Bayesian Nash equilibrium is the solution concept for static incomplete-information games:

- **2.1 Nash $\Rightarrow$ 4.3 BNE.** BNE is Nash on the agent-normal-form game.
- **4.1, 4.2 $\Rightarrow$ 4.3.** Types and type spaces are the prerequisites; BNE is the equilibrium on them.
- **4.3 $\Rightarrow$ 4.4 PBE.** For dynamic games, PBE extends BNE with belief consistency at information sets.
- **4.3 $\Rightarrow$ 4.5 Signaling games.** Signaling equilibria are refinements of BNE.
- **4.3 $\Rightarrow$ 6.1 Mechanism design.** The designer chooses rules such that BNE yields desired outcomes.
- **4.3 $\Rightarrow$ 6.3 Auction theory.** Bidders play BNE of the induced Bayesian auction game.
- **4.3 $\Rightarrow$ 9.3 CFR.** CFR’s convergence theorem is to an ϵ-BNE (or equivalent Nash) in the extensive-form representation.

Mathematical Development

Equivalence with Nash in the agent-normal-form game

Given Bayesian game G^B, define the **agent-normal-form game** G^{ANF}:

- **Players:** one per (player, type) pair (i, θ_i) — call these “agents.”
- **Strategies:** each agent (i, θ_i) picks $\sigma_i(\theta_i) \in \Delta(S_i)$.
- **Payoffs:** agent (i, θ_i)’s payoff is $U_i(\sigma_i(\theta_i) \mid \theta_i, \sigma_{-i}) \cdot p(\theta_i)$ (weighting by prior probability of being this type).

Then the Nash equilibria of G^{ANF} correspond exactly to the Bayesian Nash equilibria of G^B. Nash’s existence theorem applied to G^{ANF} gives BNE existence for finite Bayesian games.

Purification (Harsanyi, 1973)

Harsanyi showed that almost every mixed-strategy Nash equilibrium of a complete-information game is the *limit* of pure-strategy Bayesian Nash equilibria of a perturbed incomplete-information game. Specifically, add a small independent payoff perturbation $\epsilon\theta_i$ to each player’s payoffs, with θ_i drawn from a smooth distribution. The resulting Bayesian game has a pure-strategy BNE (each type plays a pure action), and as $\epsilon \to 0$, the distribution of actions induced by the pure BNE converges to the mixed strategy Nash of the original game.

This is Harsanyi’s **purification theorem**. It provides an epistemic justification for mixed strategies: the randomness is not arbitrary — it reflects hidden private information that would deterministically select an action at each type.

Efficiency and information aggregation

In Bayesian games, equilibria are generically *not* ex-post efficient. Even if the complete-information game would have a Pareto-optimal equilibrium, under incomplete information the interim rationality constraints rule out many ex-post-efficient profiles. This motivates mechanism design (Module 6): by carefully structuring the game, the designer can enforce efficient outcomes as BNE.

Myerson–Satterthwaite (1983). In a bilateral trade setting with private values, no BNE of any trading mechanism can achieve ex-post efficiency while maintaining budget balance and voluntary participation. This is a negative result: information asymmetry *forces* inefficiency in equilibrium.

Beauty-contest / level-k refinements

BNE assumes common knowledge of rationality and common prior. Empirically, subjects often play at "level-k" — optimal response to level-$k-1$ opponents. Cognitive hierarchy models (Camerer–Ho–Chong, 2004) model this as a distribution over levels. For $k \to \infty$, level-k thinking converges to BNE. For finite k, it gives a bounded-rationality refinement.

Worked Example

The Bayesian coordination game (completed)

Two players choose $\{A, B\}$. Player 1 is type-less (one type). Player 2 has $\theta_2 \in \{H, L\}$ with $p(H) = p(L) = 1/2$.

Payoffs: - Player 1 (independent of type): $u_1(A, A) = 2, u_1(B, B) = 1, u_1(\text{mismatch}) = 0$. - Player 2 type H: $u_2(A, A) = 1, u_2(B, B) = 0, u_2(\text{mismatch}) = 0$ — prefers coordinating on A. - Player 2 type L: $u_2(A, A) = 0, u_2(B, B) = 1, u_2(\text{mismatch}) = 0$ — prefers coordinating on B.

Guess: Player 2's strategy: type H plays A, type L plays B.

Player 1's expected payoff from A: $0.5 \cdot 2 + 0.5 \cdot 0 = 1$ (from type H match, type L mismatch).

Player 1's expected payoff from B: $0.5 \cdot 0 + 0.5 \cdot 1 = 0.5$ (from type H mismatch, type L match).

Player 1 plays A.

Player 2 best-response check given Player 1 plays A: - Type H gets 1 from A (coordination), 0 from B (mismatch). Plays A. ✓ - Type L gets 0 from A (mismatch), 0 from B (mismatch). Indifferent; any action is a best response.

So a BNE exists: $(A, (A, \text{anything}))$. The L-type's action doesn't matter because it doesn't affect Player 1's best response. There are multiple BNE parameterized by L-type's action.

Non-trivial case. Change Player 1's payoffs to $u_1(A, A) = u_1(B, B) = 2, u_1(\text{mismatch}) = 0$ and Player 2's weight $p(H) = 0.4, p(L) = 0.6$. Then Player 1 playing A yields $0.4 \cdot 2 + 0.6 \cdot 0 = 0.8$; playing B yields $0.4 \cdot 0 + 0.6 \cdot 2 = 1.2$. Player 1 plays B. Player 2 type H best response to Player 1 = B: type H gets 0 from A (mismatch) and 0 from B (mismatch — type H wants A). Type H is indifferent. For an equilibrium, type H plays some action, both yield 0. Check if there's an equilibrium where type H plays A: Player 1's expected payoff from A is $0.4 \cdot 2 + 0.6 \cdot 0 = 0.8$ (vs 1.2 from B); Player 1 still plays B. So (B, (A, B)) is a BNE with Player 2 effectively miscoordinating in type H and coordinating in type L.

First-price auction (two bidders, uniform values)

Bidders $i \in \{1, 2\}$ with private values $v_i \sim U[0, 1]$. Each bids $b_i \geq 0$; highest bid wins and pays bid.

Guess: symmetric linear strategy $b_i(v_i) = \alpha v_i$ for some $\alpha \in (0, 1)$.

Given opponent plays $b_j(v_j) = \alpha v_j$, bidder 1 of type v_1 bids b_1 to maximize:

$$U_1(b_1 \mid v_1) = (v_1 - b_1) \cdot \Pr[b_1 > \alpha v_2] = (v_1 - b_1) \cdot \frac{b_1}{\alpha}.$$

FOC: $\frac{\partial}{\partial b_1}\left[(v_1 - b_1)\frac{b_1}{\alpha}\right] = \frac{v_1 - 2b_1}{\alpha} = 0 \Rightarrow b_1 = v_1/2$.

So $\alpha = 1/2$: the symmetric BNE is $b_i(v_i) = v_i/2$. Each bidder shades their bid by half to trade off winning probability against surplus.

General symmetric BNE with n bidders: $b_i(v_i) = \frac{n-1}{n} v_i$. More bidders $\Rightarrow$ less shading.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Type-contingent strategy $\sigma_i(\theta_i)$	Your range strategy: what you do with each hole-card combination at each info set
Interim best response	Your best action given your specific hand and your belief about opponent's range
Interim belief $p(\theta_{-i} \mid \theta_i)$	Your posterior over opponent's hole cards given observed betting
Bayesian Nash equilibrium	**The GTO solution of the hand** — both players' range strategies are mutual best responses
Agent-normal-form game	"Each hand is a separate pseudo-agent choosing its action" — the solver's internal representation
Purification	Mixed strategies in poker (e.g., 50/50 bet/check with KQ) reflect implicit card-removal randomness

Concrete situation

The flop spot, formally. BTN c-bet flop decision on $Kd9s4c$. BTN's range has ~200 combos; BB's range has ~250 combos. BTN's strategy is a function $\sigma_{BTN} : \Theta_{BTN} \to \Delta(\{\text{bet } 33\%, \text{bet } 75\%, \text{check}\})$. BB's strategy, after observing BTN's action, is similar.

At a Bayesian Nash equilibrium: - For every combo θ_{BTN}, BTN's action at θ_{BTN} is a best response given BB's strategy. Concretely: for each combo, EV of checking must equal EV of betting 33% must equal EV of betting 75%, if all three are in the solver's output with positive probability. If only betting 33% is positive, then betting 33% must have strictly higher EV than check and 75%. - For every combo θ_{BB}, BB's action after BTN's chosen sizing is a best response given BTN's strategy.

This is the formal Bayesian Nash equilibrium condition. The solver's output is a strategy profile satisfying this at every info set, every combo, simultaneously.

CFR computes an ε-BNE. Counterfactual regret minimization iterates on regret values at each (player, info set, action) triple and converges to a strategy profile such that no combo of any player has more than ϵ counterfactual regret. This is equivalent to an ϵ-Bayesian Nash equilibrium of the poker game — the smaller ϵ, the closer to exact BNE. "GTO" in practice is ϵ-BNE with ϵ small enough (e.g., 0.1% of the pot) that exploitation gains are negligible.

Solver mixing is purification in disguise. When the solver outputs "with $A\spadesuit K\spadesuit$ on $Kd9s4c$, bet 33%: 62%, check: 38%," this is a mixed strategy. Harsanyi's purification says this mixing is the limit of pure strategies in a perturbed Bayesian game where "nearby" combos have slightly different payoffs and deterministically choose different actions. In poker, the purification is essentially achieved by card removal: $A\spadesuit K\spadesuit$ is close to $A\heartsuit K\spadesuit$ but not identical; the solver handles them separately, and the aggregate frequency of betting across all "AKs-like" combos is ~62%. The mixing you see in the output is the composition of pure strategies over these finely distinguished combos.

What this understanding buys you

GTO = Bayesian Nash equilibrium, precisely. When a poker player says "GTO says bet 33% with 62% frequency," they are reporting the strategy component of a Bayesian Nash equilibrium at a specific information set. This is not a heuristic, not an approximation, not a suggestion — it is the formal solution concept of game theory, instantiated in poker. The entire apparatus of BNE (existence, best response, interim optimality) applies.

Every combo is a best response. The key property of BNE at your info set is that *every hand in your range* is playing a best response. If the solver has you mixing with AKs, both actions (bet and check) must yield the same EV for AKs — otherwise you should only play the better one. This is the indifference principle (from 2.3) applied to type-contingent strategies. Coaches who say "AKs is indifferent between betting and checking" are stating the BNE indifference condition.

Opponent mixing is their purification, and you exploit deviations. If your opponent plays a strategy that is not BNE — say, they always bet 33% with AKs (100% instead of 62%) — then some *other* hand in their range is no longer indifferent, and you can exploit. The exploit target is the combo whose EV calculus is broken by their deviation. This is how solver-aware humans beat non-solver opponents: detect deviations from BNE, then play the best response to the deviated strategy (not to BNE).

Interim vs ex ante matters for study. When you study a spot, you look at your strategy conditional on your hand (interim). When you look at "range performance" or "how often this range wins showdown," you are doing ex ante analysis. BNE is defined at the interim level — each hand must be best-responding. Ex ante equivalence holds under common prior. If you find an ex post analysis that differs from the interim conclusion, it is because you are implicitly averaging over something that shouldn't be averaged. Always reason at the interim level when asking "should I play this way with this specific hand."

Cross-Links

- **2.1 Nash equilibrium** — the base concept extended here.
- **4.1 Bayesian games** — the underlying framework.
- **4.2 Types and type spaces** — the strategy domain.
- **4.4 Perfect Bayesian equilibrium** — dynamic refinement.
- **4.5 Signaling games** — canonical Bayesian game with receiver best response.
- **6.1, 6.3 Mechanism design, auctions** — BNE underlies both.
- **9.3 CFR** — algorithmic computation of ϵ-BNE.

Structural Compression

Invariant: A Bayesian Nash equilibrium is a type-contingent strategy profile in which every type of every player is playing a best response to the profile at their interim belief; equivalently, it is a Nash equilibrium of the agent-normal-form game where each (player, type) pair is a separate pseudo-agent weighted by its prior probability.

Mechanism: Interim expected payoff computation via Bayes' rule on the common prior, followed by type-by-type best response optimization, with the joint fixed point constituting the equilibrium.

Cross-System Echo: Bayesian Nash equilibrium is the game-theoretic counterpart of a Bayes-optimal decision rule in statistical decision theory: in both, an agent conditions on private information to compute a posterior and acts optimally against that posterior. The difference is that in BNE the "posterior" is about opponents' types rather than nature, and the opponents are themselves acting optimally against their own posteriors — a fixed point, not a one-sided optimization.

Exercises

1. Verify that the symmetric linear BNE of a first-price auction with two bidders and uniform values is $b_i(v_i) = v_i/2$.
2. Show that every finite Bayesian game has a BNE by applying Nash's existence theorem to the agent-normal-form game.
3. Construct a Bayesian coordination game where the unique BNE has both players miscoordinating in expectation.
4. In a Bayesian prisoner's dilemma where each player's cooperation cost is private (two types, H and L, with independent draws), compute the BNE. Does it depend on the distribution over types?
5. Show that BNE is equivalent to ex-ante optimality under common prior. Where does the argument break without common prior?
6. Formulate heads-up NLHE flop spot as a Bayesian game. Specify strategies, interim beliefs (accounting for blockers), and write the BNE indifference condition for one specific combo.
7. **Argue, structurally,** why Harsanyi's purification theorem provides an epistemic justification for mixed strategies. What does it say about the "arbitrariness" of mixing?

Perfect Bayesian Equilibrium

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.4

Perfect Bayesian equilibrium (PBE) is the dynamic-game analogue of Bayesian Nash equilibrium. It extends subgame perfection (3.5) to games of incomplete information by requiring, at every information set, (i) a belief about which node in the set the game is at, and (ii) a strategy that is a best response given that belief, with beliefs derived from the strategies via Bayes' rule on the equilibrium path. PBE is the standard solution concept for dynamic Bayesian games: signaling games, screening games, reputation models, multi-stage auctions. It is also, under the hood, the solution concept every poker solver is computing — every information set has a range (belief) attached, and the strategy at that info set is a best response given the range.

Formal Definition

Let Γ be an extensive-form game with information partitions $\{\mathcal{I}_i\}$.

Assessment. A pair (σ, μ) where:

- σ is a behavioral strategy profile (a probability distribution over actions at each information set for each player).
- μ is a **belief system**: for every information set I with $|I| \geq 1$, $\mu(\cdot \mid I) \in \Delta(I)$ is a probability distribution over the nodes in I, representing the acting player's belief about which node they are at.

Sequential rationality. For every information set I belonging to player i, the strategy $\sigma_i(I)$ maximizes i's expected payoff from I onward, given the belief $\mu(\cdot \mid I)$ and the continuation strategy σ. Formally:

$$\sigma_i(I) \in \arg\max_{\alpha \in \Delta(A(I))} \sum_{t \in I} \mu(t \mid I) \cdot \mathbb{E}_\sigma[u_i \mid \text{play } \alpha \text{ at } I, \text{reach } t].$$

Bayesian consistency (on-path). For every information set I reached with positive probability under σ ($\Pr_\sigma(I) > 0$):

$$\mu(t \mid I) = \frac{\Pr_\sigma(t)}{\Pr_\sigma(I)} \text{ for every } t \in I.$$

That is, on-path beliefs are derived from Bayes' rule applied to the strategy.

PBE. An assessment (σ, μ) is a **perfect Bayesian equilibrium** if it is sequentially rational and Bayesian-consistent on the equilibrium path. Off-path beliefs are not constrained by the definition, though refinements (sequential equilibrium, D1, divinity) impose additional conditions.

PBE $\subseteq$ Nash. Every PBE is a Nash equilibrium of the extensive-form game. Not every Nash is a PBE; the non-PBE Nash rely on incredible off-path play or beliefs.

SPE relationship. In perfect-information games, PBE reduces to SPE (beliefs are trivial because every info set is singleton). In imperfect-information games, PBE is strictly stronger: it requires sequential rationality at non-singleton info sets, which SPE does not.

Intuition

Subgame perfection (3.5) requires optimality at every subgame, i.e., every subtree rooted at a singleton information set. In imperfect-information games, most information sets are *not* singletons, so SPE says nothing about them. This is a hole: a Nash equilibrium can prescribe any play at a non-singleton info set, because SPE cannot rule it out.

PBE plugs this hole by requiring that players have beliefs at every information set and that their strategies are optimal given those beliefs. The beliefs are a new object — something SPE doesn't need because it never faces nontrivial info sets. Crucially, the beliefs must be derived from the strategies via Bayes' rule wherever that is well-defined (on the equilibrium path).

Why on-path only? Because off-path info sets are reached with probability zero in equilibrium, so Bayes' rule is "0/0" there. The assessment has to assign *some* belief off-path, but nothing pins down what it should be without additional refinements. Different choices of off-path beliefs can support different equilibria — this is the source of the "multiplicity" of PBE in signaling games.

The conceptual move from SPE to PBE is: **every decision requires a belief, and every belief must be consistent with strategies where possible.** This is exactly how you would think about decision making under incomplete information — you form a posterior over the unknown state and pick the optimal action given it. PBE is this principle, applied simultaneously to every player and every decision point, with a consistency requirement tying beliefs to strategies.

Structural Role

PBE is the workhorse solution concept for dynamic incomplete-information games:

- **3.5 SPE, 4.3 BNE ⇒ 4.4 PBE.** PBE synthesizes both, requiring subgame perfection-like optimality at every info set with explicit beliefs.
- **4.4 ⇒ 4.5.** Signaling games are analyzed via PBE with belief refinements.
- **4.4 ⇒ 4.6.** Sequential equilibrium refines PBE by constraining off-path beliefs via limits of totally mixed strategies.
- **4.4 ⇒ 5.5 Chain store paradox.** Reputation models are dynamic Bayesian games solved with PBE.
- **4.4 ⇒ 9.3 CFR.** CFR computes behavioral strategies with associated beliefs at each info set, converging to an ϵ-PBE of the extensive-form poker game.

Mathematical Development

Why PBE is needed: SPE's blind spot

Consider a two-stage game: Player 1 moves first (Left or Right), Player 2 does not observe the move, then Player 2 moves (Up or Down). Player 2's information set has two nodes — one after Left, one after Right — which Player 2 cannot distinguish.

Since this info set is non-singleton, no subgame is rooted there; the only subgame is the whole game. SPE reduces to Nash for this game, and a Nash equilibrium can prescribe any play at Player 2's info set without being penalized. If the non-credible play is part of Player 1's best response (because Player 1 fears it), the equilibrium stands.

PBE fixes this: Player 2 must have a belief about Left/Right at their info set, and Player 2's action must maximize expected payoff given that belief. On the equilibrium path, the belief is derived from Player 1's strategy. Any play that is not a best response given *some* belief is ruled out.

Bayesian consistency on-path

If σ is a behavioral strategy profile and I is an info set reached with positive probability, then the probability of being at node $t \in I$ conditional on reaching I is

$$\mu(t \mid I) = \frac{\pi_\sigma(t)}{\sum_{t' \in I} \pi_\sigma(t')},$$

where $\pi_\sigma(t)$ is the probability of reaching t under σ. This is Bayes' rule: condition on the observation "we are at some node in I," posterior is the normalized reach probability.

Off-path beliefs and refinements

Off the equilibrium path, $\Pr_\sigma(I) = 0$, so the Bayes' rule formula is $0/0$. PBE leaves these beliefs unconstrained, which creates multiple equilibria. Refinements constrain them:

- **Sequential equilibrium (Kreps–Wilson, 1982).** Off-path beliefs must be limits of beliefs derived from totally mixed strategies. This forces continuity with on-path reasoning.
- **Intuitive criterion (Cho–Kreps, 1987).** For signaling games, rule out off-path beliefs that assign positive probability to a sender type that would never benefit from the off-path action.
- **Divinity (Banks–Sobel, 1987).** Stronger: off-path belief must concentrate on types most likely to have deviated given their best interests.
- **D1 and D2.** Further restrictions on divinity, based on equilibrium payoffs.

Each refinement rules out some PBE, leaving a smaller equilibrium set. For applied work, sequential equilibrium is the most common.

Perfect recall and behavioral-strategy PBE

PBE is defined over behavioral strategies because perfect recall (Kuhn) makes behavioral equivalent to mixed. The belief system is similarly well-defined because at each info set the player knows what they did before (perfect recall), so Bayes' rule gives a consistent posterior over the remaining uncertainty (opponents' types and past actions).

One-shot deviation principle for PBE

As with SPE, a sequential optimality check reduces to a one-shot deviation check: an assessment (σ, μ) is sequentially rational iff no single-info-set deviation strictly improves expected payoff given the assessment. This is the standard check for PBE.

Worked Example

Signaling game: the classic job market signaling

Adapted from Spence (1973). A worker is either high-type (H, productive) or low-type (L, unproductive). The worker's type is private. Before entering the job market, the worker chooses an education level $e \in \{0, 1\}$ (0 = no school, 1 = school). Education is costly: cost $c_H < c_L$ (low types find it more costly).

After observing the education level, a firm offers a wage $w(e)$. Firms compete, so $w(e)$ equals the expected productivity given the belief about the worker's type after observing e.

Types: $\theta \in \{H, L\}, p(H) = q$.

Payoffs. Worker's payoff = wage − education cost. Firm's payoff = 0 (zero-profit competition).

Separating PBE. H chooses $e = 1$, L chooses $e = 0$; firm belief $\mu(H \mid e = 1) = 1, \mu(H \mid e = 0) = 0$; wages $w(1) = \text{productivity}(H), w(0) = \text{productivity}(L)$.

Check PBE conditions: - Bayes consistency on-path: wages match beliefs, which match strategies. ✓ - Sequential rationality: - H: education cost c_H vs wage premium $w(1) - w(0)$. H chooses school iff $w(1) - w(0) > c_H$. Assume productivity gap $> c_H$. - L: education cost c_L vs wage premium. L chooses no school iff $c_L > w(1) - w(0)$. Assume productivity gap $< c_L$. - Both conditions hold when $c_H < \text{gap} < c_L$.

Pooling PBE. Both types choose $e = 0$; firm belief $\mu(H \mid e = 0) = q$; wage $w(0) = q \cdot \text{productivity}(H) + (1 - q) \cdot \text{productivity}(L)$. For this to be PBE, off-path belief $\mu(H \mid e = 1)$ must make neither type want to deviate. If off-path belief assigns H probability 0, then $w(1) = \text{productivity}(L)$, worse than $w(0)$ for both types — no deviation. Pooling on $e = 0$ is a PBE.

Intuitive criterion rules out pooling. Cho–Kreps: if deviating to $e = 1$ is equilibrium-dominated for L (L would never prefer it even with the most generous off-path belief), then the off-path belief should assign probability 0 to L. Under that refined belief, $w(1) = \text{productivity}(H)$, and H would deviate to $e = 1$. So the pooling equilibrium is ruled out by the intuitive criterion; only the separating equilibrium survives.

Simple poker bluff-catching PBE

Two-card AKQ game (preview of Module 4.5). P1 holds A, K, or Q uniformly; P1 can bet or check. If bet, P2 can call or fold. If check, showdown. P1's bet pools or separates types based on which cards bet.

Separating PBE. P1 bets A, checks K, Q. P2's belief on seeing a bet: $\mu(A \mid \text{bet}) = 1$. Best response: fold. But then P1 of type K and Q would prefer to bluff. Contradiction; not an equilibrium.

Mixed bluffing PBE. P1 bets A always, bets Q with probability q, checks K. P2's belief on seeing a bet: $\mu(A \mid \text{bet}) = \frac{1/3}{1/3+q/3} = \frac{1}{1+q}$. P2's indifference between calling and folding requires a specific q — the bluff frequency derived from the "minimum defense frequency" logic in standard poker. The equilibrium is a PBE with this mixing; on-path beliefs are Bayes-consistent with the strategy, and sequential rationality holds at every info set.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Behavioral strategy σ	Solver output: action frequencies at each info set for each hand
Belief system μ	The "range" at each info set — distribution over opponent hole cards given the betting history
Sequential rationality	Every hand at every info set plays a best response given the range you face
Bayesian consistency on-path	Ranges are derived from the strategy via Bayes' rule (filtering hands by their action probability)
Off-path beliefs	What you believe about opponent's hand if they take an action that is zero-frequency in equilibrium
PBE	**GTO solution** — behavioral strategies with consistent ranges at every info set, each playing best response

Concrete situation

Range updating = Bayesian consistency. You are BTN, you open 2.5bb, BB calls. The flop comes $Kd9s4c$. BB checks. Now you are at an info set with a belief about BB's hand, derived from:

1. BB's preflop calling range (the range after BB's "call" action preflop).
2. BB's flop check frequency by combo (from BB's flop strategy).

The range you face is Bayesian:

$$\mu(\theta_{BB} \mid \text{BB called preflop, checked flop}) = \frac{\Pr(\text{call preflop with } \theta_{BB}) \cdot \Pr(\text{check flop with } \theta_{BB})}{\sum_{\theta'} \Pr(\text{call preflop with } \theta') \cdot \Pr(\text{check flop with } \theta')}.$$

This is exactly the Bayesian consistency condition of PBE. When you load a solver and look at "BB's check-call range on $Kd9s4c$," you are looking at the posterior μ the solver derived by applying Bayes' rule to BB's strategy.

Sequential rationality at every info set. Given this posterior, your c-bet strategy must play best response for every combo in your own range. Each of your hands (AKs, 72o, AhTh, . . .) must be either (a) betting all sizes with identical EV to checking (mixed indifference) or (b) strictly preferring one action. PBE requires this at every info set — not just the info sets you commonly reach, but every one including river spots you rarely see.

Off-path beliefs and over-betting. Suppose the solver's equilibrium strategy for BB on the river never includes raising to $3 \times$ pot (a 300% pot raise). What happens if BB does it anyway? The info set is off-path. Your belief has to come from somewhere — but PBE alone doesn't pin it down. Sequential equilibrium says the belief is the limit of trembles (what hands would most plausibly make this raise if trembling slightly), which typically puts weight on strong value hands and possibly the rare bluff, but concentrates heavily on value. Solvers handle this by imposing a tiny exploration probability at every action, guaranteeing all info sets are on-path and beliefs are well-defined — effectively computing a sequential equilibrium, not just a PBE.

Solver = PBE computation. Every modern poker solver (PioSolver, GTOWizard, Simple Postflop) computes an approximate PBE. They iterate behavioral strategies and derive beliefs from them via Bayes' rule on-path. At each iteration, the regret minimization step is a best-response improvement, pushing the strategies toward sequential rationality. At convergence, the output is an ϵ-PBE of the extensive-form poker game.

What this understanding buys you

Range = belief, always. When a poker player says "their check-call range is mostly middle pair and backdoor draws," they are describing a belief system μ at an info set. The range is not an independent object; it is the posterior derived from the opponent's strategy via Bayes' rule. Every time you update a range ("they raised, so remove the weak hands"), you are performing Bayesian consistency for PBE.

Consistency is the discipline. If your range at a node doesn't match what Bayes' rule would give you given opponent's strategy, your reasoning is inconsistent. Good poker study checks this: "if my opponent plays this way, what does my range look like here?" PBE formalizes this check as Bayesian consistency. Bad players skip it and assume ranges that don't match actual play.

Off-path actions are where humans deviate. GTO strategies are computed with tremble-based refinements, so they have beliefs even at off-path info sets. When a human opponent takes an off-path action (e.g., a bizarre overbet), a solver-trained player can still compute a best response, using the sequential-equilibrium-style belief. A pure-PBE-without-refinement player might have ambiguous beliefs and hence ambiguous best responses. This is the practical value of refinements.

Bet sizes are screens. In a signaling game, the education level screens worker type. In poker, the bet size screens your own range: "this size comes from this range." Opponents update their belief about your hand based on your bet size — a Bayesian consistency check. If you want a bet size to carry a particular belief, your actual strategy with that size must generate that belief when Bayes' rule is applied. This is why solver-derived bet sizes are non-obvious: they are chosen so that, in equilibrium, the belief they generate supports the strategy at the next info set.

Cross-Links

- **3.5 Subgame perfect equilibrium** — the perfect-information ancestor.
- **4.1 Bayesian games** — incomplete information framework.
- **4.2 Types and type spaces** — belief content at info sets.
- **4.3 Bayesian Nash equilibrium** — static counterpart.

- **4.5 Signaling games** — canonical application domain.
- **4.6 Sequential equilibrium** — refinement with consistency of off-path beliefs.
- **5.5 Chain store paradox** — reputation via PBE in repeated interaction.
- **9.3 CFR** — algorithmic computation of approximate PBE.

Structural Compression

Invariant: A perfect Bayesian equilibrium is a pair (strategy profile, belief system) such that (i) at every information set, the player's strategy maximizes expected payoff given the belief; (ii) beliefs at information sets reached with positive probability are derived from the strategy via Bayes' rule.

Mechanism: Assign a belief to every information set, require best-response optimality at each given its belief, tie on-path beliefs to strategies via Bayes' rule, and leave off-path beliefs as free parameters (possibly refined).

Cross-System Echo: PBE is the game-theoretic analogue of a Bayes-consistent filter in statistical signal processing: at each observation point, beliefs are updated via Bayes' rule, and decisions are optimal given the current belief. The innovation in game theory is that the "observations" include strategic actions of opponents who are themselves reasoning about beliefs — a mutual-consistency condition tying beliefs and strategies in a fixed point.

Exercises

1. For the job market signaling game with $q = 0.5$, productivity gap 10, $c_H = 3, c_L = 15$, verify that the separating PBE exists and satisfies all three PBE conditions.
2. Construct an extensive-form game with two Nash equilibria, exactly one of which is PBE. Identify the belief inconsistency in the non-PBE.
3. In the two-card AKQ poker example, derive the bluff frequency q such that P2 is indifferent between calling and folding.
4. Show that in a perfect-information game, PBE reduces to SPE.
5. Define the intuitive criterion formally and apply it to a three-type signaling game to rule out some PBE.
6. For a signaling game with a continuum of sender types, compute a separating PBE (e.g., monotone signaling function).
7. **Argue, structurally,** why PBE requires beliefs at every information set, not just at those reached in equilibrium. What conceptual purpose do off-path beliefs serve, and why do refinements target them?

Signaling Games

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.5

A signaling game is the canonical two-stage Bayesian game: a **sender** with private type chooses a **signal**, then a **receiver** observes only the signal (not the type) and chooses a **response**. The sender's type, the signal, and the receiver's response jointly determine both players' payoffs. Signaling games capture an enormous range of phenomena: Spence's education signaling of ability, costly advertising, warranty as quality signal, limit pricing, sovereign debt signaling. They are the first genuinely "dynamic" Bayesian games we encounter, solved via PBE (4.4) with belief refinements. In poker, every bet is a signal — you are the sender, your hand is the type, the sizing is the signal, and the opponent is the receiver updating beliefs about your range.

Formal Definition

A **signaling game** is a two-player game with the following structure:

1. **Nature** draws a type $\theta \in \Theta$ for the Sender according to common prior $p \in \Delta(\Theta)$.
2. **Sender** observes θ and chooses a message $m \in M$.
3. **Receiver** observes m (but not θ) and chooses an action $a \in A$.
4. **Payoffs**: $u_S(\theta, m, a)$ and $u_R(\theta, m, a)$.

Sender strategy. $s_S : \Theta \to \Delta(M)$, the distribution over signals by type.

Receiver strategy. $s_R : M \to \Delta(A)$, the response to each possible signal.

Belief system. $\mu : M \to \Delta(\Theta)$, the Receiver's belief about the Sender's type after observing each message.

Perfect Bayesian equilibrium of a signaling game. A triple (s_S, s_R, μ) such that:

1. **Sender optimality.** For every θ and every m in the support of $s_S(\theta)$:

$$m \in \arg\max_{m' \in M} \mathbb{E}_{a \sim s_R(m')}[u_S(\theta, m', a)].$$

2. **Receiver optimality.** For every m in the support of s_S (image of Sender's strategy):

$$s_R(m) \in \arg\max_{a \in A} \sum_{\theta} \mu(\theta \mid m) \cdot u_R(\theta, m, a).$$

3. **Bayesian consistency (on-path).** For every m reached with positive probability:

$$\mu(\theta \mid m) = \frac{p(\theta) \cdot s_S(\theta)(m)}{\sum_{\theta'} p(\theta') \cdot s_S(\theta')(m)}.$$

Classification of equilibria:

- **Separating:** distinct types send distinct messages; the Receiver learns the type exactly from the message.
- **Pooling:** all types send the same message; the Receiver learns nothing beyond the prior.
- **Semi-separating (partial pooling):** some types share messages, others separate.
- **Babbling:** messages are uninformative (equilibrium may exist with $\mu(m) = p$ for all m).

Intuition

A signaling game asks: when information is private and action is public, can the informed party credibly convey information through their actions?

The answer depends on the **single-crossing property** of payoffs: if different types have different marginal costs for sending the same message, there can be separating equilibria where each type picks a message that is cheaper for them than for other types. If types have identical cost structures, separation is impossible — every message is equally attractive across types — and only pooling or babbling survives.

Spence's education example is paradigmatic. High-productivity types can earn an education more cheaply (effort per year) than low-productivity types, so education is a "costly signal" that sorts types. In equilibrium, high types education themselves, low types do not, and the wage differential reflects the separating information.

A signaling game has *lots* of PBE because off-path beliefs are unconstrained. You can often support a pooling equilibrium by assuming "if the Sender deviated, the Receiver would think the Sender is the worst type," which deters any deviation. The **intuitive criterion** (Cho–Kreps, 1987) rules out unreasonable off-path beliefs by asking: "could this deviation benefit any type at all?" If the answer is "only type H could benefit from this deviation," then off-path belief should put probability 1 on H, not on L. Refinements shrink the equilibrium set dramatically.

For signaling to succeed (separate types), the signal must be **costly** and **differentially costly**: cheap signals can be faked by low types, so they don't separate. This is the "cheap talk" limit (Crawford–Sobel, 1982): when messages are costless, separation is much harder, and equilibria involve coarse partitioning of the type space.

Structural Role

Signaling games are the canonical two-stage Bayesian game and the simplest vehicle for PBE refinements:

- **4.1, 4.4 ⇒ 4.5.** Bayesian games + PBE → signaling games.
- **4.5 ⇒ 4.6.** Sequential equilibrium refines PBE via trembles; signaling games are where this matters most.
- **4.5 ⇒ 5.5 Chain store paradox.** Reputation is signaling across repeated interaction.
- **4.5 ⇒ 6.1 Mechanism design.** Screening (the dual of signaling) is the uninformed party designing the game; signaling is the informed party.
- **4.5 ⇒ 6.3 Auction theory.** Bid = signal of value; auction design is mechanism design for screening via bids.
- **4.5 ⇒ poker theory.** Every bet is a signal. Bet sizing conveys range information.

Mathematical Development

Separating equilibrium existence

A separating PBE in a two-type signaling game $\Theta = \{H, L\}$ with messages $M = \{m_1, m_2\}$ exists if there is an assignment $H \to m_H, L \to m_L$ with $m_H \neq m_L$ such that:

- Neither type prefers the other's message: $u_S(H, m_H, s_R(m_H)) \geq u_S(H, m_L, s_R(m_L))$ and similarly for L.
- The Receiver's optimal response given full revelation is $s_R(m_H)$ (believing $\theta = H$) and $s_R(m_L)$ (believing $\theta = L$).

This is the **single-crossing / sorting** condition: the message cost structure must differ between types in a way that makes the assignment incentive-compatible.

Pooling equilibrium existence

A pooling PBE has all types sending the same message m^*. Receiver belief on m^* is the prior: $\mu(\theta \mid m^*) = p(\theta)$. Receiver plays best response to prior.

For this to be PBE, off-path beliefs at other messages must deter deviation. Typically, $\mu(\theta \mid m') =$ some distribution that makes $s_R(m')$ sufficiently unattractive that no type wants to send m'.

The intuitive criterion

Definition. A PBE satisfies the intuitive criterion if for every off-path message m', the Receiver's belief $\mu(\theta \mid m')$ places probability zero on types θ for which

$$\max_{a \in A} u_S(\theta, m', a) < u_S(\theta, m^*, s_R(m^*))$$

where m^* is θ's equilibrium message. In words: if type θ can never benefit from sending m' (even under the most favorable Receiver response), then the Receiver should not think θ sent m'. Types that could conceivably benefit from a deviation are "candidates" for having made it.

The intuitive criterion rules out many pooling equilibria: if only the high type could benefit from an off-path education deviation, the off-path belief must concentrate on H, so the Receiver pays the H wage, so H deviates — the pooling breaks.

Cheap talk (Crawford–Sobel, 1982)

If the message is costless — $u_S(\theta, m, a)$ is independent of m — then every separating equilibrium requires that the Sender types strictly prefer their "assigned" message given the Receiver's response. Without cost differences, this is impossible for generic games: all types have the same payoff from every message-response combination, so any separating profile is indifferent.

Crawford–Sobel theorem. In a cheap-talk game where the Sender's private type is a real number $\theta \in [0, 1]$, the Sender sends a message, the Receiver chooses an action, and the Sender's bias is b (preferring higher actions by b), equilibria are finite "interval partitions": the Sender's type space is partitioned into intervals, and each interval sends the same message. The number of intervals is bounded by $1/(2b)$; for $b \geq 1/2$, only babbling equilibria exist.

The Crawford–Sobel result shows that cheap talk can convey *some* information but never full separation, and the amount of information that can be conveyed shrinks as interests diverge.

Other applications

- **Spence (1973).** Education as productivity signaling.
- **Milgrom–Roberts (1986).** Costly advertising as quality signaling (for "bad" products, advertising is not worth it at any response; "good" products advertise).
- **Kreps–Wilson, Milgrom–Roberts (1982).** Limit pricing as deterrence: incumbent monopolist signals low cost by pricing low, deterring entry.
- **Grossman (1981), Grossman–Hart (1981).** Full disclosure: under some conditions, the Sender reveals everything voluntarily because silence is interpreted as worst-case.

Worked Example

A two-type signaling game with both separating and pooling equilibria

$\Theta = \{H, L\}, p(H) = 0.5$. Messages $M = \{E, N\}$ (E = education, N = no education). Receiver chooses wage $w \in \{w_H, w_L\}$ with $w_H > w_L$. Sender payoff: $w - c(\theta, m)$ where $c(\theta, E) =$ education cost, $c(\theta, N) = 0$. Assume

$c(H,E) = 2, c(L,E) = 8$. Receiver payoff: $-(w - \text{productivity}(\theta))^2$, productivity(H) = 10, productivity(L) = 4.

Separating equilibrium attempt. $s_S(H) = E, s_S(L) = N$. Beliefs $\mu(H \mid E) = 1, \mu(H \mid N) = 0$. Receiver best response: $s_R(E) = 10, s_R(N) = 4$.

Check Sender optimality: - H: E gives $10 - 2 = 8$; N gives 4. H prefers E. ✓ - L: E gives $10 - 8 = 2$; N gives 4. L prefers N. ✓

Separating PBE exists.

Pooling equilibrium attempt. Both types send N. Belief $\mu(H \mid N) = 0.5$. Wage $s_R(N) = 0.5 \cdot 10 + 0.5 \cdot 4 = 7$. Off-path E: if belief puts probability 1 on L, wage is 4, sender payoff H: $4 - 2 = 2 < 7$, L: $4 - 8 = -4 < 7$. Neither deviates. Pooling PBE exists.

Apply intuitive criterion. For the pooling on N, consider the off-path message E. Can L benefit from E under any Receiver response? Max $u_S(L, E, a) = 10 - 8 = 2 < 7$. L can never benefit. So the belief after E should put probability 0 on L, probability 1 on H. Then Receiver pays $w_H = 10$, and H's deviation payoff is $10 - 2 = 8 > 7$ — H deviates. Pooling PBE fails the intuitive criterion.

Conclusion. Only the separating PBE survives the intuitive criterion. This is the "Spence lesson": with the single-crossing condition, refinements select the separating equilibrium.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Sender's type θ	Your hole cards
Message m	Your bet sizing and action (bet small, bet big, check, overbet, shove)
Receiver action a	Opponent's response: call, fold, raise, by what amount
Sender strategy s_S	Your bet sizing strategy: what you do with each hand
Receiver strategy s_R	Opponent's response strategy: what they do vs each of your sizings
Belief μ	Opponent's range inference from your sizing — "if you bet big, what do you have?"
Separating equilibrium	Different sizings correspond to different ranges (polarized vs merged)
Pooling equilibrium	One-size strategy where all your hands bet the same amount — opponent learns nothing from sizing
Intuitive criterion	Solver refinement: beliefs off-path make sense given "what hands would actually want to do this?"

Concrete situation

Bet sizing as signal in a river spot. You are BTN, facing BB check on the river. Your options: bet 33% pot, bet 75% pot, overbet 150% pot, or check. Each sizing is a potential message.

Pooling strategy (one size). If you always bet 75% regardless of hand, your sizing conveys no information about your hand beyond "I'm betting." BB's range-vs-your-75% is your entire betting range. This is a pooling equilibrium: all types send the same message.

Separating strategy (polarized sizing). Solver outputs often split sizings: bet small (33%) with medium-value hands, overbet (150%) with nuts/air polarized. This is a separating (actually partial-separating) equilibrium: different types send different messages, and BB's range inference differs by sizing. On seeing 33%, BB's belief is "medium value"; on seeing 150%, BB's belief is "nuts or air."

Why separation works. The payoffs must satisfy a single-crossing condition. The solver's single-crossing is the EV structure: for nuts, overbetting is cheaper (BB folds more, so you gain by applying max pressure) than betting small. For medium value, overbetting is expensive (you get called only by better, losing more). So different sizings are "cheaper" for different hand categories — the essential signaling condition.

Mixed-strategy partial separation. In practice, most solver outputs are *partial* separation: at each sizing, some range of hands is mixing. This is because perfectly separating would be detectable and counterable; the mixing creates sufficient range ambiguity that BB cannot respond deterministically. It's a semi-separating equilibrium.

Receiver best response. BB, on seeing each sizing, computes the posterior over your hand and plays best response. This is literally the PBE receiver-optimality condition. A solver enforces it at every sizing. The output: "vs your 33%, BB folds 40%, calls 60%; vs your 150%, BB folds 65%, calls 30%, raises 5%." These frequencies are optimal given the belief, which is Bayes-consistent with your sizing strategy.

What this understanding buys you

Every bet size is a message in a formal signaling game. When you choose 33% vs 75% vs overbet, you are choosing a signal. The opponent is receiving the signal and updating their belief about your range. This is not metaphor — it is the formal structure of a signaling game.

Polarization is signaling theory at work. "Polarized range" means "this sizing carries a message: 'I have nuts or nothing.' " The solver polarizes because the separation earns EV: nut hands get called less (by putting opponent in a tough spot) and air hands get fewer folds (by bluffing credibly). Both effects rely on the opponent's Bayesian inference from sizing to range.

Balance = equilibrium condition, not an aesthetic preference. "Balanced" ranges are PBE strategies: the bluff-to-value ratio at each sizing must be such that the opponent is indifferent between calling and folding. This is the receiver-optimality condition of the signaling game. If your sizing is "unbalanced," the opponent exploits by always calling (if you're too value-heavy) or always folding (if you're too bluff-heavy).

Intuitive criterion explains why certain deviations don't work. If you try to bluff with a hand that would never benefit from the bluffing sizing (e.g., medium pair overbetting — medium pair doesn't want to fold out opponent's bluff-catchers and doesn't get called by worse), the intuitive criterion says opponent should put belief 0 on that hand for that sizing. So the opponent's response deters the deviation. Solvers implement this automatically: hands only mix into sizings where they benefit.

Cheap talk doesn't work at the poker table. Poker is not cheap talk — every action has payoff consequences. This is why bet sizing *can* carry information: sending a high signal costs real chips. In pure cheap talk (just chatting), you couldn't credibly convey hand strength via language alone (beyond very coarse partitions) because there's no cost structure separating types. But the moment you back the message with money, signaling becomes possible.

Cross-Links

- **4.1 Bayesian games** — base framework.
- **4.2 Types and type spaces** — Sender's private information.
- **4.4 Perfect Bayesian equilibrium** — solution concept.
- **4.6 Sequential equilibrium** — refinement via trembles.
- **5.5 Chain store paradox** — reputation as multi-stage signaling.
- **6.1 Mechanism design** — screening (dual of signaling).

- **6.3 Auction theory** — bid as signal of value.

Structural Compression

Invariant: A signaling game is a two-stage Bayesian game in which an informed Sender sends a message and an uninformed Receiver responds, with the message partially revealing the Sender's type through Bayes-consistent beliefs; equilibrium outcomes depend on whether single-crossing conditions allow costly signaling to separate types.

Mechanism: Single-crossing cost structure enables incentive-compatible separating equilibria; pooling and babbling equilibria arise when signaling is cheap or types are too similar; refinements (intuitive criterion, divinity) use off-path belief logic to select among multiple equilibria.

Cross-System Echo: Signaling games formalize the general phenomenon of "credible revelation through costly action" — seen in evolutionary biology (handicap principle: peacock tail is a costly signal of fitness), labor economics (education as productivity signal), advertising (brand investment as quality signal), and political communication (policy commitments as type signals). In each case, only actions that are differentially costly across types can separate types credibly.

Exercises

1. Solve a two-type signaling game with messages $\{E, N\}$ and your choice of single-crossing cost. Characterize all separating and pooling PBE.
2. Show that the intuitive criterion rules out pooling equilibria in the Spence job market model when productivity gap exceeds the high type's cost.
3. In a cheap-talk game with Sender bias $b = 0.1$, apply the Crawford–Sobel analysis to determine the maximum number of intervals and thus the coarsest separation.
4. Construct a signaling game with three types and find a semi-separating equilibrium in which two types pool on one message and the third separates.
5. Show that in the Milgrom–Roberts advertising model, high-quality firms signal through costly advertising that would be unprofitable for low-quality firms.
6. Model a poker river bet-sizing decision as a signaling game with three hand types (value, medium, bluff) and two sizings (small, big). Solve for a semi-separating PBE and interpret.
7. **Argue, structurally,** why single-crossing is the key condition for separating equilibria in signaling games. What goes wrong if the single-crossing property fails?

Sequential Equilibrium

Position in Knowledge Graph

System: Game Theory **Structural Domain:** Information and Beliefs **Node:** 4.6

Sequential equilibrium (Kreps–Wilson, 1982) is the refinement of perfect Bayesian equilibrium that imposes a **consistency** condition on off-path beliefs: beliefs must be limits of beliefs derived from sequences of totally mixed strategies. This removes the "anything goes off-path" freedom of PBE and selects a tighter equilibrium set. For finite extensive-form games, sequential equilibrium is the standard solution concept — strictly stronger than PBE, weaker than refinements like proper equilibrium, and compatible with Kuhn's theorem so it is computable via behavioral strategies. Every modern poker solver effectively computes a sequential equilibrium because trembling-based exploration in iterative algorithms like CFR naturally yields consistent off-path beliefs.

Formal Definition

Let Γ be a finite extensive-form game with perfect recall. Recall an **assessment** is a pair (σ, μ) where σ is a behavioral strategy profile and μ is a belief system.

Consistency. An assessment (σ, μ) is **consistent** if there exists a sequence of *totally mixed* behavioral strategy profiles $\sigma^k \to \sigma$ such that the induced belief systems μ^k (where μ^k at every info set is derived from σ^k via Bayes' rule — well-defined because σ^k is totally mixed) satisfy $\mu^k \to \mu$.

A totally mixed strategy assigns positive probability to every action at every information set, so every node of the tree is reached with positive probability and Bayes' rule is well-defined everywhere.

Sequential rationality. As in PBE: at every information set, the strategy is a best response given the belief and the continuation play.

Sequential equilibrium. An assessment (σ, μ) is a **sequential equilibrium** if:

1. It is sequentially rational.
2. It is consistent.

Theorem (Kreps–Wilson). Every finite extensive-form game has a sequential equilibrium. Every sequential equilibrium is a PBE. Every sequential equilibrium is a Nash equilibrium. For generic finite games, sequential equilibria and PBE coincide on the equilibrium path.

Relation to trembling-hand perfection. Sequential equilibrium is implied by trembling-hand perfect equilibrium (Selten, 1975) but not equivalent. Both use trembles to discipline off-path play; sequential equilibrium requires consistency of beliefs, trembling-hand requires optimality against trembles.

Intuition

PBE leaves off-path beliefs unconstrained, which allows implausible equilibria. You can support pooling by assuming "if the Sender deviated, the Receiver would believe the Sender is the worst type." Nothing in PBE rules this out, but intuition says the belief is strange — why would the Receiver interpret a deviation as the worst news?

Sequential equilibrium disciplines off-path beliefs by asking: **if the players were trembling slightly, what beliefs would they form?** The answer is: apply Bayes' rule to the tremble distribution. As the trembles vanish, the beliefs in the limit are the "consistent" beliefs.

Crucially, the tremble sequence is *chosen by the theorist*, not unique. Different trembles give different limit beliefs. Sequential equilibrium requires the existence of *some* tremble sequence that generates the belief system, not all of them. So it is a consistency check — "are these beliefs derivable from a plausible tremble?" — rather than a unique prediction.

The practical consequence: sequential equilibrium rules out pathological PBE but often leaves multiple equilibria. Further refinements (intuitive criterion, divinity, proper equilibrium) discipline off-path beliefs more. Sequential equilibrium is the "Goldilocks" refinement: strong enough to rule out the crazy cases, weak enough to retain existence and computability.

Structural Role

Sequential equilibrium is the standard solution concept for finite extensive-form games with imperfect information:

- **3.5 SPE $\Rightarrow$ 4.4 PBE $\Rightarrow$ 4.6 Sequential equilibrium.** Strict refinement chain.
- **4.6 $\Rightarrow$ 9.3 CFR.** CFR with exploration converges to an ϵ-sequential equilibrium of the extensive-form game.
- **4.6 $\Rightarrow$ 4.5 Signaling game refinements.** Sequential equilibrium is the baseline for signaling refinements; intuitive criterion and divinity strengthen it further.
- **4.6 $\Rightarrow$ 10.1 CFR/solver theory.** Sequential equilibrium is what poker solvers compute.

Mathematical Development

Why PBE is too weak

Consider a three-node game: Sender chooses A or B, Receiver observes (in the signaling form), chooses response. Multiple PBE exist with identical on-path play but different off-path beliefs.

Example. Sender type space $\{H, L\}$, messages $\{M_1, M_2\}$, Receiver actions $\{U, D\}$. Suppose on-path play is pooling on M_1; off-path M_2 can carry any belief. PBE 1: $\mu(H \mid M_2) = 1$, PBE 2: $\mu(H \mid M_2) = 0$. Different equilibria.

Sequential equilibrium disciplines this: the belief $\mu(H \mid M_2)$ must be the limit of $\mu^k(H \mid M_2)$ where σ^k is totally mixed. If $\sigma^k(H)(M_2) = \epsilon_H^k$ and $\sigma^k(L)(M_2) = \epsilon_L^k$, then $\mu^k(H \mid M_2) = \frac{p(H)\epsilon_H^k}{p(H)\epsilon_H^k + p(L)\epsilon_L^k}$, which depends on the ratio $\epsilon_H^k / \epsilon_L^k$ as $k \to \infty$. Different ratios give different limit beliefs — so multiple sequential equilibria still exist, but the set is strictly smaller than PBE.

Consistency and the Kreps–Wilson theorem

Theorem (Kreps–Wilson, 1982). Every finite game with perfect recall has a sequential equilibrium.

Proof sketch. Define the set of consistent assessments as the set of limits of Nash equilibria of ϵ-perturbed games (games where actions are required to be played with probability at least ϵ). This set is nonempty (by existence of Nash in each perturbed game), and compactness gives the limit. The sequential rationality requirement further selects from this set.

Relationship to trembling-hand perfect equilibrium

Trembling-hand perfect (Selten, 1975). An equilibrium is trembling-hand perfect if it is the limit of Nash equilibria of totally mixed perturbations where the trembles are optimal against the unperturbed strategy.

Relationship. THP $\Rightarrow$ sequential equilibrium. Not vice versa in general extensive-form games. Sequential equilibrium is weaker: it requires belief consistency but not strategic optimality against the trembles. THP is the stronger refinement.

For two-player games with perfect recall, sequential equilibrium and THP coincide on-path but can differ off-path. In practice, sequential equilibrium is used because it is more computationally tractable.

Consistency via the "structural" constraints

For a pair (σ, μ) to be consistent, we need the existence of $\sigma^k \to \sigma$ totally mixed with $\mu^k \to \mu$. Analytically, this means the ratios $\sigma^k(I)(a)/\sigma^k(I)(a')$ can be chosen to vanish at rates matching the required belief ratios in the limit. The set of consistent beliefs can be characterized explicitly in small games but is combinatorial in general.

Sequential equilibrium and the Harsanyi common prior

If the game originated as a Bayesian game under the Harsanyi transformation, the common prior p determines the chance-node probabilities at the root. The sequence of trembles does not perturb chance; only player actions. So off-path beliefs about chance moves are still derived from p, while off-path beliefs about past player actions are determined by the trembles.

Worked Example

A small signaling game

$\Theta = \{H, L\}, p(H) = 0.6$. Messages $\{M_1, M_2\}$. Receiver actions $\{U, D\}$. Payoffs:

Type	Message	Receiver action	(Sender, Receiver)
H	M1	U	(3, 2)
H	M1	D	(0, 0)
H	M2	U	(2, 1)
H	M2	D	(0, 0)
L	M1	U	(3, 0)
L	M1	D	(0, 2)
L	M2	U	(1, 0)
L	M2	D	(0, 1)

Pooling on M_1. Receiver belief on M_1: $\mu(H \mid M_1) = 0.6$. Receiver best response: compare U giving $0.6 \cdot 2 + 0.4 \cdot 0 = 1.2$ vs D giving $0.6 \cdot 0 + 0.4 \cdot 2 = 0.8$. U wins. Receiver plays U on M_1.

Sender check: H gets 3 from M_1 (U); M_2 off-path. L gets 3 from M_1; M_2 off-path. Neither wants to deviate if off-path M_2 leads to D.

Off-path M_2: what belief? In PBE, any belief is allowed. In sequential equilibrium, the belief must be the limit of trembles.

Suppose trembles: $\sigma^k(H)(M_2) = \epsilon, \sigma^k(L)(M_2) = 2\epsilon$ (L trembles twice as often toward M_2). Then $\mu^k(H \mid M_2) = \frac{0.6\epsilon}{0.6\epsilon+0.4\cdot 2\epsilon} = \frac{0.6}{1.4} \approx 0.43$. As $\epsilon \to 0$, limit belief is 0.43. Receiver best response: U gives $0.43 \cdot 1 + 0.57 \cdot 0 = 0.43$; D gives $0.43 \cdot 0 + 0.57 \cdot 1 = 0.57$. D wins.

So with these trembles, $\mu(H \mid M_2) = 0.43, s_R(M_2) = D$. Check if the pooling is still an equilibrium: H on M_2 gets 0 (under D); H prefers M_1 (3). L on M_2 gets 0 (under D); L prefers M_1 (3). Pooling survives.

Different trembles, different sequential equilibria. Had we chosen trembles where H trembles twice as often toward M_2, $\mu(H \mid M_2) = \frac{0.6 \cdot 2}{0.6 \cdot 2 + 0.4} = 0.75$. Then receiver plays U on M_2 (payoff 0.75 vs 0.25). Sender check: H on M_2 gets 2; H on M_1 gets 3 — H still prefers M_1. L on M_2 gets 1; L on M_1 gets 3 — L still prefers M_1. Pooling still holds, but now with different off-path response.

Both are sequential equilibria — the set is genuinely multiple. Further refinements (intuitive criterion, divinity) can discipline this further.

Poker Anchor

Concept mapping

Formal object	Poker instantiation
Totally mixed strategy	Every action has positive probability at every info set — no "never do this"
Tremble	A rare alternative action taken with small probability (exploration)
Consistency	Off-path beliefs are what you'd conclude if opponent "trembled" into the off-path action
Sequential equilibrium	PBE where off-path ranges are derivable from small trembles — the standard solver output
Off-path belief	"What range does opponent have if they do something bizarre I'd never expect in GTO?"

Concrete situation

Tree exploration in CFR. When CFR iterates to approximate a Nash/sequential equilibrium, it maintains some positive probability of exploring every action at every info set during training. This prevents info sets from becoming "invisible" to the algorithm: even an action that is eventually not part of the equilibrium gets sampled occasionally, so the algorithm learns its value and updates beliefs about the reach distribution.

This is *exactly* the tremble-based definition of sequential equilibrium: the algorithm is computing Nash of an ϵ-perturbed game, and as $\epsilon \to 0$, the limit is a sequential equilibrium.

Off-path beliefs in solver output. The solver's output tells you, for every info set, the optimal action and the range of each player at that info set. Crucially, this includes info sets that are "off-path" in the equilibrium — e.g., info sets reached only via a rare mix or a dominated action. The range at those info sets is the sequential-equilibrium consistent belief, derived from the tremble sequence that CFR's exploration implicitly uses.

Consequence for human play. When a human opponent plays a line that is zero-frequency in GTO (an overbet that GTO never makes), your GTO-trained instinct is to compute a best response given the sequential-equilibrium belief — which is the range the solver has assigned to that off-path info set. Solvers handle this transparently: you can query the solver at an off-path node and it will show you the (consistent) range and the best response. Without the sequential equilibrium refinement, the belief would be undefined, and you would have no principled way to compute a best response.

Exploration vs exploitation. Sequential equilibrium captures the idea that beliefs must be "learnable" — reachable from some sequence of realized play. Trembles are exploration; consistency is learnability. When you train a poker AI via self-play, you must include exploration (random play with small probability) to keep all info sets reachable. Without this, the AI would have undefined beliefs at off-path nodes, and its strategy would be exploitable by a human taking "weird" actions. The ϵ-exploration used in training is the computational implementation of sequential equilibrium.

What this understanding buys you

Solver output is sequential equilibrium, not just PBE. Every time you load a spot into a solver, the output is an ϵ-sequential equilibrium. The off-path ranges you see are consistent (derivable from small trembles), not arbitrary. This means you can trust off-path recommendations as coherent even if the line seems bizarre — the solver has internally computed the tremble-consistent belief and played best response to it.

"I would never do that" is the opposite of sequential equilibrium. When a human says "I never bluff here, so opponent should believe I'm always strong when I bet," they are claiming an off-path belief that violates consistency. A solver-trained opponent would never take the belief at face value: they compute what the "trembled" belief would be if you had a tiny bluff frequency, and respond based on that. The solver's output will never have zero-bluff-frequency at any info set because the sequential equilibrium refinement rules out such non-trembling strategies.

Off-path surprises are handled by the refinement. Sequential equilibrium gives you a principled answer to "what do I do when opponent does something unexpected?" The answer: compute the consistent off-path belief (by imagining small trembles from equilibrium play) and best-respond to that belief. In practice, this means: treat the unexpected action as a rare but possible tremble, update accordingly, and play accordingly. You do not panic, do not assume the worst, do not assume the best. You compute.

The "intuitive criterion" is used in study. When a coach says "your 3-bet range shouldn't include this hand because it would never profit even in the best case," they are applying the intuitive criterion — a refinement of sequential equilibrium that further disciplines off-path beliefs. Sequential equilibrium alone would allow many off-path beliefs; the intuitive criterion selects those consistent with incentive-compatibility. Both together give the tightest principled answer to "what belief should I have here."

Cross-Links

- **3.5 Subgame perfect equilibrium** — the perfect-information refinement.
- **4.4 Perfect Bayesian equilibrium** — the PBE that this refines.
- **4.5 Signaling games** — the main application domain.
- **4.3 Bayesian Nash equilibrium** — the static Bayesian counterpart.
- **9.3 CFR** — the computational algorithm that approximates sequential equilibrium.
- **10.1 Game theory of poker** — sequential equilibrium is the formal solution concept.

Structural Compression

Invariant: A sequential equilibrium is a perfect Bayesian equilibrium whose beliefs are additionally required to be consistent — expressible as the limit of beliefs generated by totally mixed behavioral strategies — so that off-path beliefs cannot be arbitrary but must derive from some plausible tremble sequence.

Mechanism: ϵ-perturbation of the strategy space, Bayes' rule applied within each perturbation to derive consistent beliefs, and limit-taking as $\epsilon \to 0$ to characterize the refined equilibrium set.

Cross-System Echo: Sequential equilibrium is the game-theoretic analogue of "learnability under exploration" in reinforcement learning. An RL agent that explores all actions with positive probability learns a value function on all states, including those rarely visited; without exploration, off-path values are undefined and the policy is brittle. Sequential equilibrium is the fixed point of "play optimally assuming all states can be visited by trembles," which is the same logic as ε-greedy exploration.

Exercises

1. Show that every sequential equilibrium is a PBE but not every PBE is a sequential equilibrium. Give a concrete game where these differ.

2. For a three-stage signaling game with two sender types and three messages, compute a sequential equilibrium by specifying a consistent tremble sequence.

3. Verify that in a perfect-information game, sequential equilibrium coincides with SPE.

4. Show that sequential equilibrium exists for every finite game with perfect recall (sketch the Kreps–Wilson proof).
5. Construct a game where different trembles yield different sequential equilibria with distinct off-path beliefs.
6. In a poker river decision with an unexpected overbet, formulate the decision as an off-path info set and compute the sequential equilibrium belief derived from a plausible tremble sequence.
7. **Argue, structurally,** why sequential equilibrium is the natural "learning" refinement for extensive-form games. How does the tremble-based definition connect to the computational practice of exploration in self-play algorithms like CFR?