Perfect Bayesian Equilibrium

Game Theory · Module 4 — Information And Beliefs

Node 4.4

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?

Game Theory · Module 4 — Information And Beliefs · Node 4.4
Concepts: bayesian-updating, bayesian-nash-equilibrium, conditional-probability, nash-equilibrium
Prerequisites: 3.5, 4.3
Enables: 4.5, 4.6, 5.5

hbar.university — own.your.cognition