Sequential Equilibrium

Game Theory · Module 4 — Information And Beliefs

Node 4.6

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 ⇒ 4.4 PBE ⇒ 4.6 Sequential equilibrium.** Strict refinement chain.
- **4.6 ⇒ 9.3 CFR.** CFR with exploration converges to an ϵ-sequential equilibrium of the extensive-form game.
- **4.6 ⇒ 4.5 Signaling game refinements.** Sequential equilibrium is the baseline for signaling refinements; intuitive criterion and divinity strengthen it further.
- **4.6 ⇒ 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?

Game Theory · Module 4 — Information And Beliefs · Node 4.6
Concepts: bayesian-nash-equilibrium, nash-equilibrium, conditional-probability
Prerequisites: 4.4
Enables: 9.3

hbar.university — own.your.cognition