

Lecture 11: Equilibrium in IIEFGs: Beliefs, Seq. Rationality, and PBE

Lecturer: Swaprava Nath

Scribe(s): Claude Scribe

Disclaimer: These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.

11.1 Recap and the Question for Today

Lecture 9 and 10 gave us the machinery to *describe* play under imperfect information: information sets, behavioral strategies, perfect recall, and Kuhn’s theorem telling us that with perfect recall the mixed and behavioral descriptions are interchangeable.

What we do not yet have is a notion of *stability*. In a PIEFG we ruled out non-credible threats using subgame perfection: restrict the profile to each subgame and demand a Nash equilibrium there. That device relies on being able to say “the play is now at node x , start again from here”. Under imperfect information a player often cannot say which node she is at, so the device needs repair.

Two ingredients replace it. First, a **belief**: a probability distribution over the nodes of an information set. Second, **sequential rationality**: at each information set, the player’s action maximises expected utility *against her belief*. Together they give the **Perfect Bayesian Equilibrium**.

11.2 Where Subgame Perfection Still Works, and Where It Stops

Example 11.1 (Release or Review) A lab (Player 1) has a model ready. It can exit the process (E) and release under the existing licence, or send the model for external review (X). If it goes to review, the lab ships either the tuned build (T) or the base build (B). The auditor (Player 2) receives a stripped artefact, cannot tell the two builds apart, and chooses a light audit (L) or a rigorous one (R). The game is in Figure 11.1.

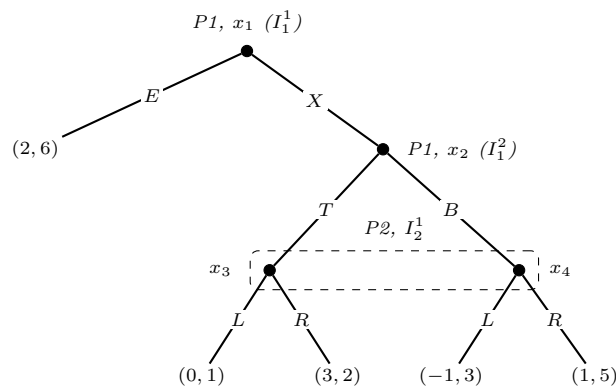


Figure 11.1: The Release-or-Review game of Example 11.1. Payoffs are (Lab, Auditor).

This game has exactly two subgames, the whole game at x_1 and the one rooted at x_2 . There is no subgame rooted at x_3 , because x_3 and x_4 belong to the same information set of Player 2. A subgame is required not to cut an information set: were we to start a “game” at x_3 , we would be handing the auditor the information that the build is the tuned one, which she does not have.

Backward induction still goes through here, because I_2^1 lies entirely inside the subgame at x_2 . In that subgame the auditor’s payoffs are 1 from L and 2 from R at x_3 , and 3 from L and 5 from R at x_4 ; R strictly dominates L , so she audits rigorously whatever she believes. Given R , the lab compares 3 from T against 1 from B and picks T , so the subgame is worth (3, 2). At the root, X gives 3 against 2 from E , so the lab sends the model for review.

Remark 11.2 *We were lucky. Because R dominated L pointwise, the auditor did not need a belief at all. In general an information set will straddle the boundary of every candidate subgame, no domination will be available, and subgame perfection will have nothing to say. That is the situation the rest of this lecture handles.*

11.3 Beliefs and Sequential Rationality

Definition 11.3 (Belief) *Let the information sets of player i be $I_i = \{I_i^1, I_i^2, \dots, I_i^{k(i)}\}$. A **belief** of player i is a collection of mappings $\mu_i^j : I_i^j \rightarrow [0, 1]$ with $\sum_{x \in I_i^j} \mu_i^j(x) = 1$ for each j . It is the conditional distribution over histories in the information set, conditioned on having reached that information set.*

A belief is a free object: any distribution over the nodes is a belief. The next definition ties it to the strategy profile actually being played.

Definition 11.4 (Bayesian belief) *Write $P_\sigma(x)$ for the probability that the play visits vertex x when the players use the behavioral profile σ . A belief $\mu_i = \{\mu_i^1, \dots, \mu_i^{k(i)}\}$ is **Bayesian with respect to σ** if it is derived from σ by Bayes’ rule,*

$$\mu_i^j(x) = \frac{P_\sigma(x)}{\sum_{y \in I_i^j} P_\sigma(y)}, \quad \forall x \in I_i^j, \forall j = 1, \dots, k(i). \quad (11.1)$$

Definition 11.5 (Sequential rationality) *A strategy σ_i of player i at an information set I_i^j is **sequentially rational** given σ_{-i} and the partial belief μ_i^j if*

$$\sum_{x \in I_i^j} \mu_i^j(x) u_i(\sigma_i, \sigma_{-i} \mid x) \geq \sum_{x \in I_i^j} \mu_i^j(x) u_i(\sigma'_i, \sigma_{-i} \mid x) \quad \text{for every } \sigma'_i. \quad (11.2)$$

The pair (σ, μ) is called an **assessment**, and it is sequentially rational if (11.2) holds for every player at every information set.

In plain words (see Remark 11.7 below for a fuller picture): a belief is what a player thinks she is facing inside an information set; sequential rationality just says she picks the action with the best expected payoff given that belief, nothing more. She is not asked to guess correctly every time, only to act optimally given what she believes.

11.4 A Worked Example

Example 11.6 (A three-round agent duel; MSZ Example 7.38) An agent (Player 1) plays against an adaptive opponent (Player 2). The agent opens with one of three probes L, M, R ; the opponent responds with ℓ, m, r without seeing which probe was used; the agent then acts again from a position it can only partly identify. The game has perfect recall, so mixed and behavioral strategies are equivalent. It is drawn in Figure 11.2.

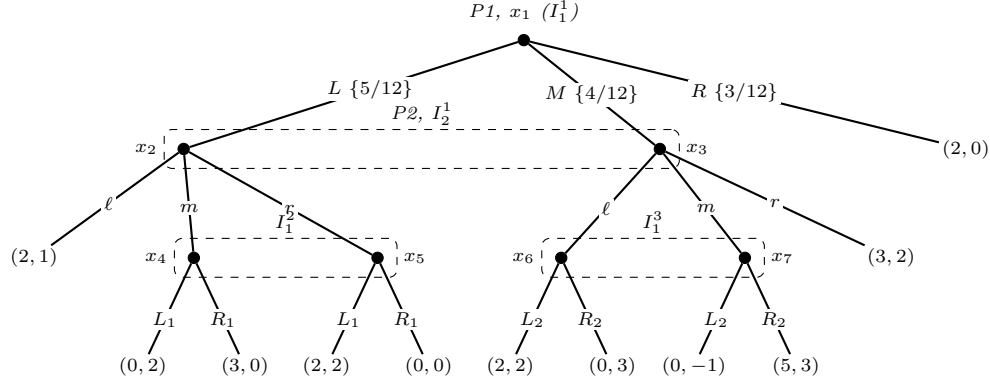


Figure 11.2: The agent duel of Example 11.6. Payoffs are (Agent, Opponent). Player 1 owns $I_1^1 = \{x_1\}$, $I_1^2 = \{x_4, x_5\}$ and $I_1^3 = \{x_6, x_7\}$; Player 2 owns $I_2^1 = \{x_2, x_3\}$.

Consider the behavioral profile σ given by

$$\begin{aligned}\sigma_1(I_1^1) &= (L\{5/12\}, M\{4/12\}, R\{3/12\}), & \sigma_2(I_2^1) &= (\ell\{1\}, m\{0\}, r\{0\}), \\ \sigma_1(I_1^2) &= (L_1\{0\}, R_1\{1\}), & \sigma_1(I_1^3) &= (L_2\{1\}, R_2\{0\}).\end{aligned}$$

Step 1: the reaching probabilities. Reading off the tree,

$$P_\sigma(x_2) = \frac{5}{12}, \quad P_\sigma(x_3) = \frac{4}{12}, \quad P_\sigma(x_4) = 0, \quad P_\sigma(x_5) = 0, \quad P_\sigma(x_6) = \frac{4}{12}, \quad P_\sigma(x_7) = 0. \quad (11.3)$$

Nodes x_4 and x_5 are reached only if Player 2 plays m or r at x_2 , which σ_2 never does; x_7 is reached only through m at x_3 , likewise never.

Step 2: Player 2's belief, and whether ℓ is right. Applying (11.1) at I_2^1 ,

$$\mu_2^1(x_3) = \frac{P_\sigma(x_3)}{P_\sigma(x_2) + P_\sigma(x_3)} = \frac{4/12}{5/12 + 4/12} = \frac{4}{9}, \quad \mu_2^1(x_2) = \frac{5}{9}.$$

Under σ , playing ℓ at x_2 terminates at $(2, 1)$, and playing ℓ at x_3 leads to x_6 , where Player 1 takes L_2 and the play ends at $(2, 2)$. So the opponent's expected payoff from ℓ is

$$\frac{5}{9} \times 1 + \frac{4}{9} \times 2 = \frac{13}{9}, \quad (11.4)$$

which is larger than what m or r deliver. So ℓ is sequentially rational for Player 2 given μ_2^1 .

Step 3: Player 1 at I_1^3 , and the threshold. Let $p = \mu_1^3(x_7)$. Reading the four leaves under I_1^3 ,

$$u_1(L_2) = (1 - p) \cdot 2 + p \cdot 0 = 2 - 2p, \quad u_1(R_2) = (1 - p) \cdot 0 + p \cdot 5 = 5p.$$

Hence R_2 is strictly better exactly when $5p > 2 - 2p$, that is when

$$p > \frac{2}{7}. \quad (11.5)$$

Under σ we have $P_\sigma(x_6) = 4/12$ and $P_\sigma(x_7) = 0$, so $p = 0$ and L_2 is correct. Had the belief put more than $2/7$ on x_7 , Player 1 should have switched to R_2 . The action is neither right nor wrong on its own; it is right or wrong relative to a belief.

Step 4: Player 1 at I_1^1 . Playing R ends at $(2, 0)$. Playing L leads to x_2 , where Player 2 plays ℓ and the play ends at $(2, 1)$. Playing M leads to x_3 , then ℓ to x_6 , then L_2 to $(2, 2)$. All three give Player 1 utility 2, so she is indifferent and any mixture, including $(5/12, 4/12, 3/12)$, is sequentially rational.

Every player is sequentially rational at every information set against a belief derived from σ . The assessment is therefore an equilibrium in the sense we are about to define.

Remark 11.7 (A simple analogy: the doctor behind the curtain) *It helps to picture a doctor who is told: “your next patient has either flu or COVID, and you will only be told the symptoms, not which one it is.” She cannot see the patient’s file (the information set), only the room she has been sent into. Before meeting the patient, the doctor already knows how common each illness is this season – that background rate is her belief. If she is given no other information, her belief is exactly last season’s flu-vs-COVID ratio: how often patients like this actually turned out to have one or the other. That is precisely what Bayes’ rule computes in (11.1): whenever a room can genuinely be visited, the belief about what’s in it must match how often patients like this actually end up there. In our tree, $\mu_2^1(x_2) = 5/9$ and $\mu_2^1(x_3) = 4/9$ are exactly this kind of ratio, read off from how often the Agent’s strategy sends the Opponent to x_2 versus x_3 .*

Remark 11.8 (The 0/0 problem: a room that is never visited) *Now suppose a particular room is one the clinic’s booking system never actually sends anyone to. What should the doctor believe if she ever finds herself there anyway? There is no historical ratio to consult: nobody has ever been in that room, so “how often does flu vs. COVID end up here” simply has no answer. In our tree this is exactly $I_1^2 = \{x_4, x_5\}$: we have $P_\sigma(x_4) = P_\sigma(x_5) = 0$, so the ratio in (11.1) is 0/0, and Definition 11.4 imposes no restriction at all. Off the equilibrium path, any belief is admissible, because there is no data to be consistent or inconsistent with. This freedom is not a defect of our bookkeeping; it is a genuine gap in the solution concept, and disciplining it is exactly what the sequential equilibrium of MSZ Chapter 7 does. It is also why $\sigma_1(I_1^2) = R_1$ went unchecked above: whatever belief we put on $\{x_4, x_5\}$, we could have made R_1 optimal.*

11.5 Perfect Bayesian Equilibrium

Definition 11.9 (Perfect Bayesian Equilibrium) *An assessment (σ, μ) is a **Perfect Bayesian Equilibrium (PBE)** if for every player $i \in N$,*

1. μ_i is Bayesian with respect to σ , and
2. σ_i is sequentially rational given σ_{-i} and μ_i .

A PBE is often quoted by naming only σ , since on the equilibrium path μ is determined by σ . It is a refinement of Nash equilibrium, and it is self-enforcing in the same sense as SPNE: no player wants to deviate even at an information set she does not expect to reach.

The following theorem says the new notion is a genuine generalisation of the old one.

Theorem 11.10 *In a PIEFG, a behavioral strategy profile σ is an SPNE if and only if $(\sigma, \hat{\mu})$ is sequentially rational, where $\hat{\mu}$ is the degenerate belief.*

Proof:[Sketch] In a PIEFG every information set is a singleton $\{x\}$, so the only distribution on it is $\hat{\mu}(x) = 1$; there is nothing for a belief to encode. Condition (11.2) then reads $u_i(\sigma_i, \sigma_{-i} \mid x) \geq u_i(\sigma'_i, \sigma_{-i} \mid x)$ for every x , which is precisely the requirement that the restriction of σ to the subgame rooted at x be a Nash equilibrium of that subgame, for every x . That is the definition of SPNE. In the doctor analogy of Remark 11.7, a PIEFG is a clinic where the doctor is always handed the patient's file: there is no room with two possible patients in it, so there is nothing left for a belief to do. ■

Remark 11.11 (Why this is the right abstraction for learning agents) *Maintaining a distribution over hidden states and best-responding to it is exactly the belief-state formulation of a POMDP, and the same object appears in opponent modelling for poker, security games and negotiation agents. Remark 11.8 has a practical counterpart too: a learned policy is never trained on states its own behaviour never visits, so both Bayes' rule and gradient descent leave the same hole, and both get exploited through it.*

11.6 Game Theory in Practice: P2P File Sharing

The rest of the lecture is a case study in what happens when the off-path behaviour of a protocol is left undefended. The material follows Parkes and Seuken, Chapter 5 [PS17].

11.6.1 Desired properties and terminology

A peer-to-peer system wants **scalability** and **failure resilience**. The vocabulary: a **protocol** is the set of messages that can be sent and actions that can be taken over the network; a **client** is a particular process for sending those messages; a **reference client** is a particular implementation; a **peer** is a participant.

11.6.2 Early systems

Napster (1999–2001) used a centralized database and let users download music from each other. **Gnutella** (2000–) removed the server: a peer obtains a list of IP addresses from known peers, broadcasts a query, receives a query response routed back from any peer holding the file, and then downloads directly.

Gnutella's problem is visible as soon as one writes down the one-shot game between two peers, Table 11.1.

Free-riding strictly dominates sharing: $3 > 2$ against Share, and $0 > -1$ against Free-ride. The unique PSNE is (Free-ride, Free-ride) with payoff $(0, 0)$, Pareto dominated by $(2, 2)$. This is a Prisoner's Dilemma, and the prediction held: about 85% of peers were free-riding by 2005, and Gnutella accounted for less than 1% of worldwide P2P traffic by 2013. Client developers could have enforced sharing, but they compete with one another, and a client that forces you to upload loses users to one that does not.

		Peer 2	
		Share	Free-ride
Peer 1	Share	2, 2	-1, 3
	Free-ride	3, -1	0, 0

Table 11.1: The (Gnutella) file sharing game.

11.6.3 BitTorrent

BitTorrent (2001–) fixed this with one structural idea: *break the file into pieces*. A single exchange becomes a repeated game, and “if you let me download, I will reciprocate” becomes enforceable. The leecher-seeder interaction is a repeated Prisoner’s Dilemma and the seeder’s strategy is tit-for-tat.

The **tracker** is a centralized entity that tracks connections between peers and their upload and download speeds. The reference client protocol is:

- Set a threshold r of uploading speed, typically the third-highest speed observed in the recent past.
- If a peer j uploaded to i at a rate $\geq r$, unchoke j in the next period.
- If a peer j uploaded to i at a rate $< r$, choke j in the next period.
- Every three time periods, **optimistically unchoke** a random currently-choked peer from the neighbourhood, and leave it unchoked for three periods.

Optimistic unchoking exists because tit-for-tat has a bootstrapping problem: a newcomer has nothing to trade, so somebody must move first. It is the protocol’s deliberate off-path generosity, and it is also its attack surface.

11.6.4 Strategic behaviour and attacks

A peer chooses how often to contact the tracker, which pieces to reveal, how many upload slots to open and at what speed, and what data to allow others to download. Possible goals range from minimising upload to maximising download speed.

BitThief [Loc06] aims to download without ever uploading. It keeps asking the tracker for peers so as to grow its neighbourhood as fast as possible, and lives entirely off the optimistic unchoke slots it receives from that large neighbourhood. It never uploads. The fix is at the tracker: block repeat requests from the same IP address within 30 minutes.

The strategic piece revealer [Lev08] attacks the piece-selection rule instead. The reference client tells neighbours about new pieces and requests using “rarest-first”. The manipulator instead reveals the *most common* piece that a reciprocating peer does not have, protecting a monopoly on the rare pieces and keeping others interested in trading with it.

BitTyrant manipulates the choice of how much to upload to whom, allocating bandwidth to the peers that give the best marginal return rather than following the reference rule.

11.6.5 Takeaways

P2P demonstrates the importance of game theory in computer systems. Early systems were easily manipulated. BitTorrent's innovation was to break files into pieces, which enabled tit-for-tat. Vulnerabilities remain, but it is on balance a very successful example of incentive-based protocol design.

11.7 Next Time

Lecture 12 changes the source of uncertainty. Here players were uncertain about *where in a known game* they stood. Next we take players who are uncertain about *which game they are playing at all*: games with incomplete information, types, common priors, and Bayesian equilibrium.

- [MSZ] M. Maschler, E. Solan, S. Zamir, *Game Theory*, Cambridge University Press. Chapters 6 and 7.
- [PS17] D. C. Parkes and S. Seuken, *Economics and Computation*, Chapter 5.
- [Loc06] T. Locher, P. Moor, S. Schmid, R. Wattenhofer, "Free Riding in BitTorrent is Cheap", HotNets 2006.
- [Lev08] D. Levin, K. LaCurts, N. Spring, B. Bhattacharjee, "BitTorrent is an Auction: Analyzing and Improving BitTorrent's Incentives", SIGCOMM 2008.