

Game Theory 14.122: Handout #1

Finding PBE in Signaling Games

1 General Strategy

In a 2×2 signaling game, there can be any or all of the following Perfect Bayesian Equilibria (PBE): both types of Player 1 may play pure strategies in equilibrium (if they play the same strategy, we say it is a *pooling equilibrium*; if they differ, we say it is a *separating equilibrium*); one type of Player 1 may play a pure strategy while the other plays a mixed strategy (leading to a semi-separating equilibrium); or both types of Player 1 may play mixed strategies. (We won't deal with the latter case.)

When looking for a PBE...

1. Decide whether you're looking for a separating, pooling, or semi-separating equilibrium.
2. Assign a strategy (a message for each type) to Player 1; make sure it is not strictly dominated.
3. Derive beliefs for Player 2 according to Bayes' rule at each information set reached with positive probability along the equilibrium path. Set arbitrary beliefs for Player 2 at information sets that are never reached along the equilibrium path.
4. Determine Player 2's best response.
5. In view of Player 2's response, check to see whether Player 1 has an incentive to deviate from the strategy you assigned her in any state of *the* world (in other words, for all types of Player 1). If she does not, you have found a PBE. If she does, this is not an equilibrium return to step 2 and assign Player 1 a different strategy.
6. Once you have exhausted all possible strategies within an equilibrium subset, return to step 1 and select a different type of PBE.

2 Example

Nature moves first. With equal probabilities, Nature assigns Player 1 type $t \in \{t_1, t_2\}$.

Player 1 knows her type, and chooses whether to play L or R.

Player 2 does not know what player's 1 type is, but he sees whether she plays L or R. He chooses U or D.

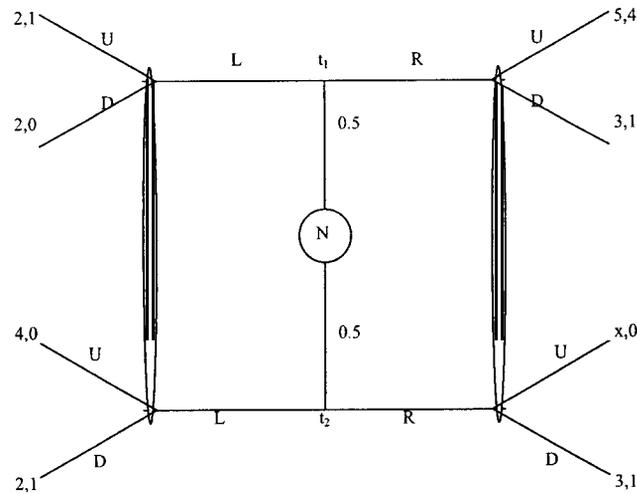


Figure 1: Game

The extensive form of the game and its payoffs are presented in Figure 1. For example, if Player 1 is of type t_1 and plays R, and Player 2 plays U, the payoffs are 5 to Player 1 and 4 to Player 2.

Note that if Player 1 is of type t_2 and plays R, and Player 2 plays U, the payoff to Player 1 is x . Different values of x will be used at different points.

Note also that for Player 1 of type t_1 , R strictly dominates L: playing L, she will always get 2; playing R, she will get at least 3. This means that in any PBE, type t_1 will always play R. This is intended to simplify the search for equilibria by eliminating half of the possible equilibria of each type (more than half of the semi-separating equilibrium possibilities). In a more general set up, these additional possibilities would need to be explored.

3 Finding a Separating PBE

3.1 Player 1 Strategy

Assume for now that $x = 2$.

In a separating PBE, each Player 1 type chooses a different message, so that the message perfectly identifies the player type. There are only two possibilities for separating strategies for Player 1. One possibility is that type 1 will choose R and type 2 will choose L. The other possibility is that type 1 will choose L and type 2 will choose R. If either type has one message which strictly dominates the other, that type of Player 1 will never play the dominated strategy, so the selection of strategies to test is simplified. In this case, because for type t_1 , R strictly dominates L as explained above if there exists a separating PBE, it must be that type 1 plays R and type 2 plays L. In other words,

$$\sigma_1(t) = \begin{cases} R & \text{if } t = t_1 \\ L & \text{if } t = t_2 \end{cases}.$$

3.2 Beliefs for Player 2

Let $\mu(t_i|A)$ be the probability Player 2 assigns to type i after observing action A .

If Player 2 sees that Player 1 plays L (in other words, in the left-hand information set) she will assign probability 1 to type 2. If she sees Player 1 play R, she will assign probability 1 to type 1. These are the only beliefs consistent with Bayes' rule, because both the left-hand and the right-hand information sets are reached with positive probability along the equilibrium path.

To see this, recall Bayes' rule:

$$\mu(t_1|R) = \frac{P(R|t_1)P(t_1)}{P(R)} = \frac{P(R|t_1)P(t_1)}{P(R|t_1)P(t_1) + P(R|t_2)P(t_2)}$$

We know from the structure of the game that

$P(t_1) = P(t_2) = 0.5$. By construction (the assignment of Player 1's strategy above) we know that

$$P(R|t_1) = 1$$

and

$$P(R|t_2) = 0.$$

Plugging in and solving, we get

$$\mu(t_1|R) = 1$$

(and therefore, $\mu(t_2|R) = 0$).

In the same way, we solve for $\mu(t_1|L) = 0$.

3.3 Best Response for Player 2

3.3.1 Against R

Player 2's expected utility from playing U and D at the right-hand information set, respectively, is

$$\begin{aligned} \mathbb{E}U_2(U, R) &= \mu(t_1|R) * U_2(U, R; t_1) + \mu(t_2|R) * U_2(U, R; t_2) = 4 \\ &\text{and} \\ \mathbb{E}U_2(D, R) &= \mu(t_1|R) * U_2(D, R; t_1) + \mu(t_2|R) * U_2(D, R; t_2) = 1. \\ \text{Therefore } BR_2(R) &= U. \end{aligned}$$

3.3.2 Against L

$$\begin{aligned} \mathbb{E}U_2(U, L) &= \mu(t_1|L) * U_2(U, L; t_1) + \mu(t_2|L) * U_2(U, L; t_2) = 0. \\ \mathbb{E}U_2(D, L) &= \mu(t_1|L) * U_2(D, L; t_1) + \mu(t_2|L) * U_2(D, L; t_2) = 1. \\ \text{Therefore } BR_2(L) &= D. \end{aligned}$$

3.4 Is this an Equilibrium?

Since Player 2's beliefs are Bayesian by construction, and Player 2's strategy is a best response given those beliefs, this is an equilibrium if and only if neither type of Player 1 has an incentive to deviate.

We already know that type t_1 will not deviate, because for her R strictly dominates L.

Type t_2 will follow the assigned strategy as long as the payoff it yields is at least as high as the one she would get if she deviated.

Type t_2 's payoff along the equilibrium path is $U_1(L, D; t_2) = 2$.

If she deviated and played R instead, Player 2's beliefs would continue to be as above; in other words, he would believe, upon seeing R played, that Player 1 was of type t_1 with probability 1, and would therefore play U. Player 1's payoff from deviating would therefore be $U_1(R, U; t_2) = x = 2$. Since this is no higher than her equilibrium payoff, she will not deviate.

Since neither type of Player 1 will deviate, the beliefs derived above are correct, and the following is a PBE:

$$\begin{aligned} \sigma_1(t) &= \begin{cases} R & \text{if } t = t_1 \\ L & \text{if } t = t_2 \end{cases} \\ \sigma_2(a_1, \mu(a_1)) &= \begin{cases} U & \text{if } a_1 = R \\ D & \text{if } a_1 = L \end{cases} \\ \text{and} \\ \mu(a_1) &= \begin{pmatrix} \mu(t_1|R) \\ \mu(t_1|L) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Note that the beliefs are an integral part of the equilibrium; these beliefs support the equilibrium strategies. If Player 2 had different beliefs, his optimal strategy might be different.

3.5 Comments

1. If type t_1 did not have a dominant strategy, we would have to also check whether a separating equilibrium in which type t_1 plays L and t_2 plays R.

Exercise 1 Suppose $x = 4$. *Show that no separating equilibrium exists.*

4 Finding a Pooling PBE

4.1 Player 1 Strategy

Assume, as before, that $x = 2$.

In a pooling PBE, both types of Player 1 choose the same message, so that they cannot be distinguished on the basis of their behavior. There are only two possibilities for separating strategies for Player 1: play R always, or play L always. In this case, because for type t_1 , R strictly dominates L as explained above if there exists a pooling PBE, both types of Player 1 must play R:

$$\sigma_1(t) = \begin{cases} R & \text{if } t = t_1 \\ R & \text{if } t = t_2 \end{cases}.$$

4.2 Beliefs for Player 2

If Player 2 sees that Player 1 plays R – along the equilibrium path – Bayes' rule applies:

$$\mu(t_1|R) = \frac{P(R|t_1)P(t_1)}{P(R)} = \frac{P(R|t_1)P(t_1)}{P(R|t_1)P(t_1) + P(R|t_2)P(t_2)}$$

We know from the structure of the game that $P(t_1) = P(t_2) = 0.5$. By construction (the assignment of Player 1's strategy above) we know that

$$P(R|t_1) = 1$$

and

$$P(R|t_2) = 1.$$

Plugging in and solving, we get

$$\mu(t_1|R) = \mu(t_2|R) = 0.5.$$

If, however, Player 1 plays L, which never occurs along the equilibrium path, Bayes' rule does not apply:

$$\mu(t_1|L) \neq \frac{P(L|t_1)P(t_1)}{P(L)}$$

because $P(L) = 0$. In this case, we get to arbitrarily assignment beliefs.

It is not possible at this point to know what, if any, beliefs regarding off-equilibrium path behavior will sustain a pooling equilibrium. So for now assume that $\mu(t_1|L) = \lambda \in [0, 1]$.

4.3 Best Response for Player 2

4.3.1 Against R

Player 2's expected utility from playing U at the right-hand information set is

$$\mathbb{E}U_2(U, R) = \mu(t_1|R) * U_2(U, R; t_1) + \mu(t_2|R) * U_2(U, R; t_2) = 2.$$

Similarly,

$$\mathbb{E}U_2(D, R) = \mu(t_1|R) * U_2(D, R; t_1) + \mu(t_2|R) * U_2(D, R; t_2) = 1.$$

Therefore U dominates D, and Player 2 will always respond to R with U:

$$BR_2(R) = U.$$

4.3.2 Against L

$$\mathbb{E}U_2(U, L) = \mu(t_1|L) \cdot U_2(U, L; t_1) + \mu(t_2|L) \cdot U_2(U, L; t_2) = \lambda.$$

$$\mathbb{E}U_2(D, L) = \mu(t_1|L) \cdot U_2(D, L; t_1) + \mu(t_2|L) \cdot U_2(D, L; t_2) = 1 - \lambda.$$

Which is the best response therefore depends on the beliefs of Player 2:

$$BR_2(L) = \begin{cases} U & \text{if } \lambda > 0.5 \\ D & \text{if } \lambda < 0.5 \\ qU + (1-q)D & \text{if } \lambda = 0.5 \quad (q \in [0, 1]) \end{cases}.$$

4.4 Is this an Equilibrium?

Since Player 2's beliefs are Bayesian by construction, and Player 2's strategy is a best response given those beliefs, this is an equilibrium if and only if neither type of Player 1 has an incentive to deviate.

We already know that type t_1 will not deviate, because for her R strictly dominates L.

Type t_2 will follow the assigned strategy as long as the payoff it yields is at least as high as the one she would get if she deviated.

Type t_2 's payoff along the equilibrium path is $U_1(R, U; t_2) = x = 2$. If she deviated and played L instead, Player 2's off-equilibrium path beliefs would kick in.

If $\lambda < 0.5$, he would play D, and Player 1's payoff from deviating would therefore be $U_1(L, D; t_2) = 2$. Since this is no higher than her equilibrium payoff, she will not deviate. This would therefore be an equilibrium.

If $\lambda > 0.5$, he would play U, and Player 1's payoff from deviating would therefore be $U_1(L, U; t_2) = 4$, which is higher than her equilibrium payoff. She will deviate, which means the strategies reassigned her will not hold up as equilibrium strategies under the specified beliefs.

If $\lambda = 0.5$, Player 2 is indifferent between U and D, and may play any (possibly degenerate) combination of the two. Because the expected payoff to Player 1 from any non-degenerate mixture by player 2 is strictly greater than 2 (the equilibrium payoff she gets), the belief $\lambda = 0.5$ can be sustained in equilibrium only if q , the probability that Player 2 plays U, is 0.

Summarizing, the following are pooling PBEs of this game:

$$\sigma_1(t) = \begin{cases} R & \text{if } t = t_1 \\ R & \text{if } t = t_2 \end{cases}$$

$$\sigma_2(a_1, \mu(a_1)) = \begin{cases} U & \text{if } a_1 = R \\ D & \text{if } a_1 = L \end{cases}$$

and

$$\mu(a_1) = \begin{pmatrix} \mu(t_1|R) \\ \mu(t_1|L) \end{pmatrix} = \begin{pmatrix} 0.5 \\ \lambda \end{pmatrix}, 0 \leq \lambda \leq 0.5.$$

Although the equilibrium strategies are uniquely defined, the above represents a continuum of equilibria in which λ takes on values along $[0, 0.5]$.

4.5 Comments

1. This example serves to illustrate the importance of off-equilibrium path beliefs. If $\lambda > 0.5$, Player 2's off-equilibrium path behavior would change. Because in equilibrium Player 1 knows Player 2's strategy (both on and off the equilibrium path), this change in Player 2's strategy would change Player 1's best response, and the equilibrium reconstructed would crumble.
2. If type t_1 did not have a dominant strategy, we would have to also check whether a pooling equilibrium in which both types of Player 1 play L exists.

Exercise 2 For what, if any, values of x does no pooling equilibrium exist?

5 Finding a Semi-Separating PBE

5.1 Player 1 Strategy

Continue to assume that $x = 2$.

In a semi-separating PBE, one type of Player 1 plays a pure strategy while the other plays a mixed strategy. As a result, Player 2 is able to imperfectly update his prior beliefs about Player 1's type. Contrast this with the case of a pooling equilibrium, in which no updating is possible (the posterior belief along the equilibrium path are the same as the priors), and with a separating equilibrium, in which updating is perfect (in each information set correct beliefs place probability 1 on one type and probability 0 on the other). For example, suppose all smart types go to MIT, while dumb types may go to either Harvard or MIT. Then if you observe someone going to Harvard, you know she's a dumb type, but if you observe someone going to MIT you cannot tell whether he is smart or dumb.

In a 2×2 (2 players, 2 pure strategies for each) game of the type with which we're concerned there are four possible semi-separating equilibria. If Player 1 of type t_1 mixes L and R, Player 1 of type t_2 could potentially play only R or only L. Similarly, if it is type t_2 which is mixing, type t_1 could play either L all of the time or R all of the time. Again, having R strictly dominate L for type t_1 greatly simplifies things in our example, because she can neither make nor play L only. Therefore, if there is a semi-separating equilibrium in this game, it must take the form that Player 1 of type t_1 plays R exclusively, and Player 1 of type t_2 mixes L and R. The only uncertainty regards the mixing probabilities for type t_2 . For now, to be completely general, let the strategy be

$$\sigma_1(t) = \begin{cases} R & \text{if } t = t_1 \\ pL + (1-p)R & \text{if } t = t_2 \end{cases}, p \in (0, 1].$$

5.2 Beliefs for Player 2

Because both L and R are played with positive probability along the equilibrium path, Bayes' rule applies in both information sets for Player 2.

$$\begin{aligned}\mu(t_1|R) &= \frac{P(R|t_1)P(t_1)}{P(R)} = \frac{P(R|t_1)P(t_1)}{P(R|t_1)P(t_1)+P(R|t_2)P(t_2)} = \frac{0.5}{0.5+0.5(1-p)} = \frac{1}{2-p} \\ \mu(t_1|L) &= 0.\end{aligned}$$

5.3 Best Response for Player 2

5.3.1 Against R

$$\begin{aligned}\mathbb{E}U_2(U, R) &= \mu(t_1|R) * U_2(U, R; t_1) + \mu(t_2|R) * U_2(U, R; t_2) = \frac{4}{2-p} \\ \mathbb{E}U_2(D, R) &= \mu(t_1|R) * U_2(D, R; t_1) + \mu(t_2|R) * U_2(D, R; t_2) = 1 \\ \text{Therefore } BR_2(R) &= U.\end{aligned}$$

5.3.2 Against L

$$\begin{aligned}\mathbb{E}U_2(U, L) &= \mu(t_1|L) * U_2(U, L; t_1) + \mu(t_2|L) * U_2(U, L; t_2) = 0 \\ \mathbb{E}U_2(D, L) &= \mu(t_1|L) * U_2(D, L; t_1) + \mu(t_2|L) * U_2(D, L; t_2) = 1 \\ \text{Therefore } BR_2(L) &= D.\end{aligned}$$

5.4 Is this an Equilibrium?

Since Player 2's beliefs are Bayesian by construction, and Player 2's strategy is a best response given those beliefs, this is an equilibrium if and only if neither type of Player 1 has an incentive to deviate.

We already know that type t_1 will not deviate, because for her R strictly dominates L.

Type t_2 will follow the assigned strategy as long as she is exactly indifferent between playing L and playing R. If she is not, she will deviate and always play the strategy which yields the highest payoff.

Type t_2 's payoff along the equilibrium path is $U_1(R, U; t_2) = x = 2$ if she plays R, and $U_1(L, D; t_2) = 2$ if she plays L. So she is indifferent, which means she is willing to play a mixture strategy, so Player 2's beliefs are correct.

Summarizing, the following are semi-separating PBEs of this game:

$$\begin{aligned}\sigma_1(t) &= \begin{cases} R & \text{if } t = t_1 \\ pL + (1-p)R & \text{if } t = t_2 \end{cases}, P \in (0, 1) \\ \sigma_2(a_1, \mu(a_1)) &= \begin{cases} U & \text{if } a_1 = R \\ D & \text{if } a_1 = L \end{cases} \\ \text{and} \\ \mu(a_1) &= \begin{pmatrix} \mu(t_1|R) \\ \mu(t_1|L) \end{pmatrix} = \begin{pmatrix} \frac{1}{2-p} \\ 0 \end{pmatrix}.\end{aligned}$$

5.5 Comments

1. Notice that in this case, because the semi-separating equilibria are supported by (almost) the entire unit interval, the separating equilibrium is

simply the limit of the semi-separating equilibria as $p \rightarrow 1$. The pooling equilibria are *not* obtained by taking the limit as $p \rightarrow 0$ because Bayes' rule fails to apply in that case.

2. If type t_1 did not have a dominant strategy, we would have to also check whether any (or all) of the other three possible configurations of a semi-separating equilibrium work in this case.

Exercise 3 Suppose $U_2(U, R; t_1) = 1.5$ (instead of $\frac{1}{4}$ as in the diagram). What (if any) restrictions would this put on the semi-separating equilibrium?

Exercise 4 Suppose $x = 1$. For what, if any, values of $U_1(U, R; t_1)$ is there a PBE in which both types of Player 1 play mixed strategies? Note that if $U_1(U, R; t_1) \leq 2$ then Player 1 of type t_1 no longer has a strictly dominant strategy.