Game Theory:
Perfect Equilibria in Extensive Form Games

Branislav L. Slantchev
Department of Political Science, University of California – San Diego

May 15, 2008

Contents.

---

1 Nash Equilibrium in Extensive Form Games

We already know how to solve strategic form games and now we also know how to convert extensive form to strategic form as well. The solution concept we now define ignores the sequential nature of the extensive form and treats strategies as choices to be made by players before all play begins (i.e. just like in strategic games).

DEFINITION 1. A **Nash equilibrium of a finite extensive-form game** Γ is a Nash equilibrium of the reduced normal form game G derived from Γ.

We can do this because the finite extensive form game has a finite strategic form. More generally though, a Nash equilibrium of an extensive form game is a strategy profile (s_i^*, s_{-i}^*) such that $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$ for each player i and all $s_i \in S_i$. That is, the definition of Nash equilibrium is the same as for strategic games (but be careful how you specify the strategies here).

Finding the Nash equilibria of extensive form games thus boils down to finding Nash equilibria of their reduced normal form representations. We have already done this with Myerson's card game, reproduced in Fig. 1 (p. 2).

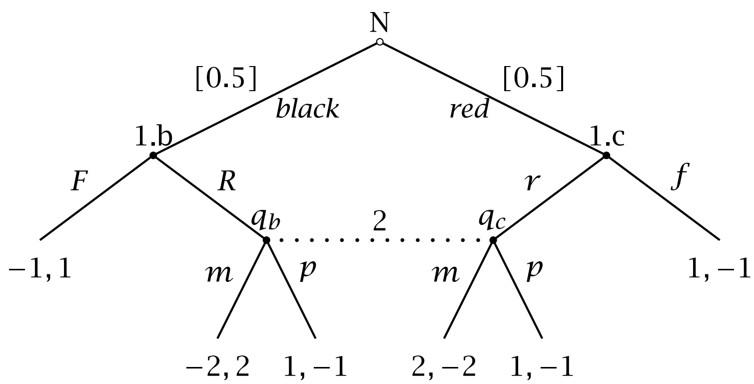


Figure 1: Myerson's Card Game in Extensive Form.

Recall that the mixed strategy Nash equilibrium of this game is:

$$\left\langle \left(\frac{1}{3}[Rr], \frac{2}{3}[Fr]\right), \left(\frac{2}{3}[m], \frac{1}{3}[p]\right) \right\rangle.$$

If we want to express this in terms of behavior strategies, we would need to specify the probability distributions for the information sets. Player 1 has two information sets, b following the black card, and c following the red card. The probability distributions are $(2/3[F], 1/3[R])$ at information set b, and $(0[f], 1[r])$ at information set c. In other words, if player 1 sees the black (losing) card, he folds with probability $2/3$. If he sees the red (winning) card, he always raises. Player 2's behavior strategy is specified above (she has only one information set).

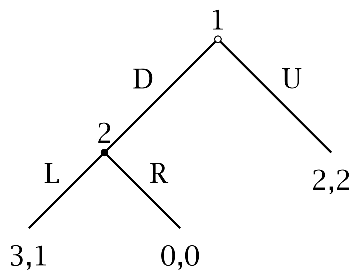
Because in games of perfect recall mixed and behavior strategies are equivalent (Kuhn's Theorem), we can conclude that a Nash equilibrium in behavior strategies must always exist in these games. This follows directly from Nash's Theorem. Hence, we have the following important result:

THEOREM 1. *For any extensive-form game* Γ *with perfect recall, a Nash equilibrium in behavior strategies exists.*

Generally, the first step to solving an extensive-form game is to find all of its Nash equilibria. The theorem tells us at least one such equilibrium will exist. We furthermore know that if we find the Nash equilibria of the reduced normal form representation, we would find all equilibria for the extensive form. Hence, the usual procedure is to convert the extensive-form game to strategic form, and find its equilibria.

1.1 Selten's Game

However, some of these equilibria would have important drawbacks because they ignore the dynamic nature of the extensive-form. This should not be surprising: after all, we obtained the strategic form representation by removing the element of timing of moves completely. Reinchard Selten was the first to argue that some Nash equilibria are "more reasonable" than others in his 1965 article. He used the example in Fig. 2 (p. 3) to motivate the discussion, and so will we.



	L	R
U	2,2	2,2
D	3,1	0,0

Figure 2: Selten's Example.

The strategic form representation has two pure-strategy Nash equilibria, $\langle D, L\rangle$ and $\langle U, R\rangle$.[1] Look closely at the Nash equilibrium (U, R) and what it implies for the extensive form. In the profile (U, R), player 2's information set is never reached, and she loses nothing by playing R there. But there is something "wrong" with this equilibrium: *if* player 2's information set is ever reached, then she would be strictly better off by choosing L instead of R. In effect, player 2 is threatening player 1 with an action that would not be in her own interest to carry out. Now player 2 does this in order to induce player 1 to choose U at the initial node thereby yielding her the highest payoff of 2. But this threat is not credible because given the chance, player 2 will always play L, and therefore this is how player 1 would expect her to play if he chooses D. Consequently, player 1 would choose D and player 2 would choose L, which of course is the other Nash equilibrium $\langle D, L\rangle$.

The Nash equilibrium $\langle U, R\rangle$ is not plausible because it relies on an incredible threat (that is, it relies on an action which would not be in the interest of the player to carry out). In fact, none of the MSNE will be plausible for that very reason either. According to our motivation for studying extensive form games, we are interested in sequencing of moves presumably because players get to reassess their plans of actions in light of past moves by other players

[1]What about mixed strategies? Suppose player 1 randomizes, in which case player 2's best response is L. But if this is the case, player 1 would be unwilling to randomize and would choose D instead. So it cannot be the case that player 1 mixes in equilibrium. What if player 2 mixes? Let q denote the probability of choosing L. Player 1's expected payoff from U is then $2q + 2(1 - q) = 2$, and his expected payoff from D is $3q$. He would choose U if $2 \geq 3q$, or $^2/_3 \geq q$, otherwise he would choose D. Player 2 cannot mix with $1 > q > 2/3$ in equilibrium because she has a unique best response to D. Therefore, she must be mixing with $0 \leq q \leq 2/3$. For any such q, player 1 would play U. So, there is a continuum of mixed-strategy Nash equilibria, where player 1 chooses U, and player 2 mixes with probability $q \leq 2/3$. These have the same problem as $\langle U, R\rangle$.

(and themselves). That is, nonterminal histories represent points at which such reassessment may occur. The only acceptable solution should be the PSNE $\langle D, L\rangle$.

The following definition is very important for the discussion that follows. It helps distinguish between actions that would be taken if the equilibrium strategies are implemented and those that should not.

DEFINITION 2. Given any behavior strategy profile σ, and information set is said to be **on the path of play** if, and only if, the information set is reached with positive probability according to σ. If σ is an equilibrium strategy profile, then we refer to the *equilibrium path of play*.

To anticipate a bit of what follows, the problem with the $\langle U, R\rangle$ solution is that it specifies the incredible action at an information set that is off the equilibrium path of play. Player 2's information set is never reached if player 1 chooses U. Consequently, Nash equilibrium cannot pin down the optimality of the action at that information set. The problem will not extend to strategy profiles which visit all information sets with positive probability. The reason for this is that if the Nash equilibrium profile reaches all information sets with positive probability, then it will also reach all outcomes with positive probability. But if it does so, the fact that no player can profit by deviating from his Nash strategy implies that there would exist no information set where he would want to deviate. In other words, his actions at all information sets are credible. If, on the other hand, the Nash strategies leave some information sets off the path of play, then the Nash requirement has no bite: whatever the player does at these information sets is "irrelevant" as it cannot affect his payoffs. It is under these circumstances that he may be picking an action that he would not never choose if the information set is actually reached. Notice that unlike $\langle U, R\rangle$, the other PSNE $\langle D, L\rangle$ does reach all information sets with positive probability. In this case, Nash's requirement is sufficient to establish optimality of the strategies everywhere. As we shall see, our solutions will always be Nash equilibria. It's just that not all Nash equilibria will be reasonable.

We now look at two examples that demonstrate that this problem occurs not only in games of certainty, complete and perfect information, but also in games of certainty with imperfect information, and games of uncertainty with imperfect information.

1.2 The Little Horsey

Consider the following simple game in Fig. 3 (p. 4). Player 1 gets to choose between U, M, or D. If he chooses D, the game ends. If he chooses either U or M, player 2 gets to choose between L and R without knowing what action player 1 has taken except that it was not D.

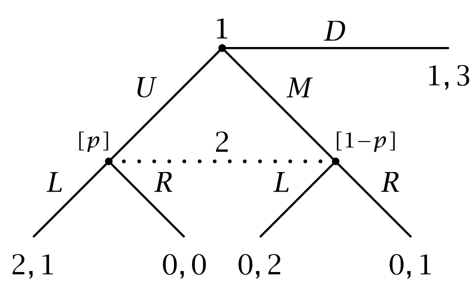


Figure 3: The Little Horsey Game.

What are the Nash equilibria? Let's convert this to normal form, the result is in Fig. 4 (p. 5). By inspection, we see that there are two Nash equilibria in pure strategies, $\langle U, L\rangle$ and $\langle D, R\rangle$.

However, there is something unsatisfying about the second one. In the $\langle D, R\rangle$ equilibrium, player 2 seems to behave irrationally. If her information set is ever reached, playing L strictly dominates R. So player 1 should not be induced to player D by the incredible threat to play L. However since player 2's information set is off the equilibrium path, Nash equilibrium does not evaluate the optimality of play there.

	L	R
U	2,1	0,0
M	0,2	0,1
D	1,3	1,3

Figure 4: The Normal Form of the Game in Fig. 3 (p. 4).

1.3 Giving Gifts

There are two players and player 1 receives a book which, with probability p is a small game theory pocket reference, and with probability $1 - p$ is a Star Trek data manual. The player sees the book, wraps it up, and decides whether to offer it to player 2 as a gift. Player 2 hates Star Trek and is currently suffering in a graduate game theory course, so she would prefer to get the game theory references but not the Star Trek manual. Unfortunately, she cannot know what is being offered until she accepts it.

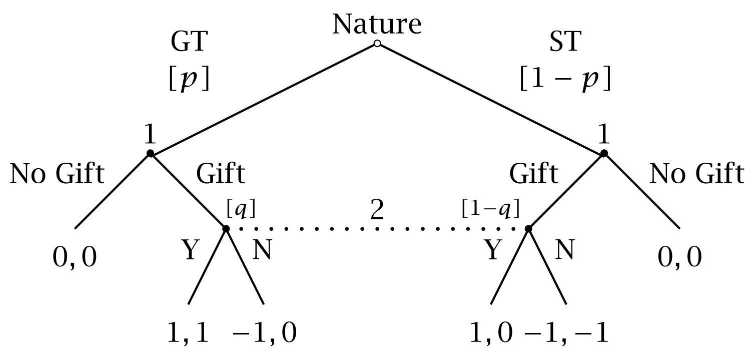


Figure 5: The Gift-Giving Game.

Consider the extensive-form game in Fig. 5 (p. 5). Player 1, who knows what gift he has for player 2, offers the wrapped gift to player 2. If the gift is accepted, then player 1 derives a positive payoff because everyone likes when their gifts are accepted. Player 1 hates the humiliation of having a gift rejected, so the payoff is -1.

Player 2 strictly prefers accepting the game theory book to not accepting it; she is indifferent between not accepting this book and accepting the Star Trek manual, but hates rejecting the Star Trek manual more than the game theory book because while dissing game theory is cool, dissing Star Trek is embarrassing.

Let's construct the strategic form of this game. Player 2 has only two strategies: accept or reject. Player 1 has two information sets, and therefore four pure strategies that specify what to do with each book. Each strategy has two components: $a_G a_S$, with a_G specifying what to do if the book is the game theory reference, and a_S specifying what to do if it is the Star Trek manual. Fig. 6 (p. 6) shows the strategic form.

	Y	N
GG	$1, p$	$-1, p-1$
GN	p, p	$-p, 0$
NG	$1-p, 0$	$p-1, p-1$
NN	$0, 0$	$0, 0$

Figure 6: Strategic Form of the Game in Fig. 5 (p. 5).

$\langle GG, Y\rangle$ is a Nash equilibrium for any value of p because $p - 1 < 0 < p < 1$. Player 1 offers the gift regardless of its type, and player 2 accepts always. In addition, $\langle NN, N\rangle$ is a Nash equilibrium. Player 1 never offers any gifts, and player 2 refuses any gifts if offered.

However, the problem with profile $\langle NN, N\rangle$ is that it prescribes an action for player 2 that is clearly irrational: *if* the game ever reaches player 2's information set, then accepting a gift strictly dominates not accepting a gift regardless of what the gift is.

Because the strategic form ignores timing, Nash equilibrium only ensures optimality at the start of the game. That is, equilibrium strategies are optimal if the other players follow their equilibrium strategies. But we cannot see whether the strategies continue to be optimal once the game begins. We now turn to solving extensive form games directly in behavior strategies. (Recall that we shall refer to them as mixed strategies.)

2 Perfect Bayesian Equilibrium

2.1 Conditional Beliefs

In the example game in Fig. 3 (p. 4), player 2 does not observe player 1's choice: if she gets to move, all she knows is that player 1 has either chosen U or M. In the game in Fig. 5 (p. 5), player 2 only observes a gift offer but does not know what the gift is. We shall require that player 2 form *beliefs* about the probability of being at any particular node in her information set. Obviously, if the information set consists of a single node, then, if that information set is reached, the probability of being at that node is 1. If there are more than one nodes in the information set, the belief will be a probability distribution over the nodes.

Let's look at the game in Fig. 5 (p. 5). Let q denote player 2's belief that she is at the left node in her information set (given that the set is reached, this is the probability of player 1 having offered the game theory book), and $1 - q$ be the probability that she is at the right node (given that the set is reached, this is the probability of player 1 having offered the Star Trek book). That is, p is player 2's **initial belief** (or the prior) of the book being a game theory reference; and q is player 2's **conditional belief** (updated belief, or posterior) about the book being a game theory reference given that player 1 has offered it as a gift.

For example, suppose that player 2 believes player 1 is playing the strategy NG. If player 2 observes a gift offer, then she concludes that the probability of the gift being the Star Trek book is 1 because only player 1 would ever offer this book as a gift. More generally, player 2 will update her belief conditional on arriving at the information set. Of course, if player 2 thinks player 1 plays the strategy GG, then the updated belief will equal the prior: Player 2 does not learn anything about the gift from the player 1's behavior. Putting all this together yields the first requirement:

REQUIREMENT 1 (Beliefs). At each information set, the player who gets to move must have a **belief** about which node in the information set has been reached by the play of the game.

Note that this requirement does not tell us anything about how these beliefs are formed (but you can already probably guess that they involve Bayes rule). Before studying this, let's see how simply having these beliefs can help solving some of the problems we encountered. For this, we need to introduce a second requirement.

2.2 Sequential Rationality

A strategy is **sequentially rational** for player i at the information set h if player i would actually want to chose the action prescribed by the strategy if h is reached. Because we require players to have beliefs for each information set, we can "split" each information set and evaluate optimal behavior at all of its nodes, including ones in information sets that are not reached along the equilibrium path. For now, we are not interested how these beliefs are formed, just that players have them. For example, regardless of what player 1 does, player 2 will have *some* updated belief q at her information set.[2]

A **continuation game** refers to the information set and all nodes that follow from that information set. Using their beliefs, players can calculate the expected payoffs from continuation games. In general, the optimal action at an information set may depend on which node in that set the play has reached. To calculate the expected payoff of an action at information set h, we have to consider the continuation game.

More formally, take any two nodes $x \succ y$ (that is, y follows x in the game tree), and consider the mixed strategy profile σ. Let $P(y|\sigma, x)$ denote the probability of reaching y starting from x and following the moves prescribed by σ. That is, $P(y|\sigma, x)$ is the conditional probability that the path of play would go through y after x if all players chose according to σ and the game started at x. It is the multiplicative product of all chance probabilities and move probabilities given by σ for the branches connecting x and y. Naturally, if y does not follow x, then $P(y|\sigma, x) = 0$. Player i's expected utility in the continuation game starting at node x then is:

$$U_i(\sigma|x) = \sum_{z \in \mathcal{Z}} P(z|\sigma, x) u_i(z),$$

where $\mathcal{Z}$ is the set of terminal nodes in the game, and $u_i(\cdot)$ is the Bernoulli payoff function specifying player i's utility from the outcome $z \in \mathcal{Z}$. Note that this payoff only depends on the components of σ that are applied to nodes that follow x. Anything prior to x is ignored. These expected utilities are therefore conditional on node x having been reached by the path of play.

For example, consider Selten's game from Fig. 2 (p. 3) and suppose $\sigma_1 = (1/3, 2/3)$, and $\sigma_2 = (0.5, 0.5)$. The continuation game from player 1's information set includes the entire game tree. What is player 1's expected payoff? Let z_1 denote the outcome (D, L), z_2 denote the outcome (D, R), and z_3 denote the outcome (U). The expected payoff then is:

$$\begin{aligned}
U_1(\sigma|\emptyset) &= P(z_1|\sigma, \emptyset)u_1(z_1) + P(z_2|\sigma, \emptyset)u_1(z_2) + P(z_3|\sigma, \emptyset)u_1(z_3) \\
&= \sigma_1(D)\sigma_2(L)(3) + \sigma_1(D)\sigma_2(R)(0) + \sigma_1(U)(2) \\
&= 1/3 \times 1/2 \times 3 + 2/3 \times 2 = 11/6.
\end{aligned}$$

[2]This is true even in the case where player 1's strategy is NN, in which case q would reflect the belief about player 1 given the unexpected, zero-probability, event that a gift is offered.

What about his payoffs in the continuation game starting at player 2's information set? It would be:

$$\begin{aligned} U_1(\sigma|D) &= P(z_1|\sigma, D)u_1(z_1) + P(z_2|\sigma, D)u_1(z_2) \\ &= \sigma_2(L)(3) + \sigma_2(R)(0) \\ &= {}^1\!/_2 \times 3 = {}^3\!/_2. \end{aligned}$$

We can also calculate player 2's payoff in both continuation games in an analogous manner:

$$\begin{aligned} U_2(\sigma|\emptyset) &= P(z_1|\sigma, \emptyset)u_2(z_1) + P(z_2|\sigma, \emptyset)u_2(z_2) + P(z_3|\sigma, \emptyset)u_2(z_3) \\ &= \sigma_1(D)\sigma_2(L)(1) + \sigma_1(D)\sigma_2(R)(0) + \sigma_1(U)(2) \\ &= {}^1\!/_3 \times {}^1\!/_2 \times 1 + {}^2\!/_3 \times 2 = {}^3\!/_2 \end{aligned}$$

and

$$\begin{aligned} U_2(\sigma|D) &= P(z_1|\sigma, D)u_2(z_1) + P(z_2|\sigma, D)u_2(z_2) \\ &= \sigma_2(L)(1) + \sigma_2(R)(0) \\ &= {}^1\!/_2 \times 1 = {}^1\!/_2. \end{aligned}$$

How does player 2 evaluate the optimality of her behavior using her beliefs? She calculates expected utilities as usual and then chooses the strategy that yields the highest expected payoff in the continuation game. That is, *if the information set h is a singleton*, the behavior strategy profile σ is sequentially rational for i at h if, and only if, player i's strategy σ_i maximized player i's expected payoff at h if all other players behaved according to σ.

In Selten's example, player 2 has only one sequentially rational strategy at her information set D, which is trivial to verify. The expected payoff in the continuation game following this node is maximized if she chooses L with certainty. For player 1, the calculation involves player 2's strategy. Player 1 would choose D if:

$$U_1(D, \sigma_2) \geq U_1(U, \sigma_2) \Leftrightarrow \sigma_2(L)(3) + \sigma_2(R)(0) \geq 2 \Leftrightarrow \sigma_2(L) \geq {}^2\!/_3.$$

Thus, strategy D is sequentially rational for any $\sigma_2(L) > 2/3$; strategy U is sequentially rational for any $\sigma_2(L) < 2/3$; and both D and U are sequentially rational if $\sigma_2(L) = 2/3$.

If information sets are singletons, sequential rationality is straightforward to define in the way we just did. What do we do for information sets that contain more than one node? This is where beliefs come in: we calculate the expected payoff from choosing a particular action conditional on these beliefs. That is, σ_i is sequentially rational for player i at the (non-singleton) information set h if, and only if, σ_i maximizes player i's expected payoff when node $x \in h$ occurs given the belief probabilities that player i assigns to the nodes x given that h has occurred, and assuming that the play continues according to σ.

The intuition is as follows. When a player finds himself at some information set, he must be able to assess the consequences of choosing some action. This is straightforward when the information set is a singleton: the player will know precisely what will happen given his choice. But what if the information set is not a singleton? In that case, the same action could have different consequences depending on which node in the information set the player is actually at. This is where beliefs come in. Since a belief is just a probability distribution over the nodes in this information set, the player can compute the expected payoff from choosing a particular action. To do this, he takes the payoff from taking that action at some node in the information set and multiplies it by the probability that he is at that node; he does that for all nodes in the information set and adds the results together. In other words, the usual expected payoff calculation!

For example, consider the game in Fig. 3 (p. 4), and let p denote player 2's belief that she is at the left node in her information set conditional on this set having been reached. Her expected payoff from choosing L is then $p \times (1) + (1-p) \times (2) = 2-p$; analogously, her expected payoff from choosing R is $p \times (0) + (1-p) \times (1) = 1-p$. She would choose L whenever $2-p > 1-p$; that is, always (this inequality holds regardless of the value of p). Hence, the unique sequentially rational strategy for player 2 at her information set is to choose L with certainty.

A similar thing happens with the gift game from Fig. 5 (p. 5). Player 2 can calculate the expected payoff from choosing Y, which is $q(1)+(1-q)(0) = q$, and the expected payoff from choosing N, which is $q(0)+(1-q)(-1) = q-1$. Since $q > q-1$ for all values of q, it is never optimal for player 2 to choose N regardless of the beliefs player 2 might hold. Therefore, the strategy N is not sequentially rational because there is no belief that player 2 can have that will make it optimal at her information set. In other words, the unique sequentially rational strategy is to choose Y with certainty.

We now require that in equilibrium players should only choose sequentially rational strategies (actually, we can prove that this must be the case given beliefs, but we need to learn how to derive these beliefs first).

Requirement 2 (Sequential Rationality). Given their beliefs, the players' strategies must be **sequentially rational**. That is, at each information set the actions by the players must form a Nash equilibrium in the continuation game.

This is now sufficient to rule out the implausible Nash equilibria in the three examples we have seen. Consider Selten's game. As we saw above, player 2's unique sequentially rational strategy is L (and it forms a Nash equilibrium in the trivial decision game at her information set). Therefore, $\sigma_2^*(L) = 1$ in equilibrium. Since choosing D is sequentially rational for player 1 for any $\sigma_2(L) > 2/3$, this implies that D is the unique equilibrium sequentially rational rational strategy for player 1. We conclude that the only Nash equilibrium that involves sequentially rational strategies is $\langle D, L\rangle$. In other words, we ruled out all implausible equilibria.

Let's apply this to the game in Fig. 3 (p. 4). As we have seen, the expected utility from playing L then is $2-p$ and the expected utility from playing R is $1-p$. Since $2-p > 1-p$ for all values of p, Requirement 2 prevents player 2 from choosing R. Simply requiring player 2 to have beliefs and act optimally given these beliefs eliminates the implausible equilibrium $\langle D, R\rangle$.

Finally, as we have seen in the game from Fig. 5 (p. 5), the only sequentially rational strategy for player 2 is Y. Hence, Requirement 2 eliminates the implausible equilibrium $\langle NN, N\rangle$ there as well.

Let's now look at a more interesting example in Fig. 7 (p. 10). This is the same game as the one in Fig. 5 (p. 5) but with a slight variation in the payoffs. This actually looks a lot more like a gift-giving game. In this situation, player 2 strictly prefers not to accept the Star Trek book, and is indifferent between rejecting gifts regardless of what they are. Accepting the preferred gift is the best outcome. Player 1, in turn, still wants his gift accepted but he also prefers to please player 2. Even then, he'd rather have her accept even the gift she does not like than suffer the humiliation of rejection.

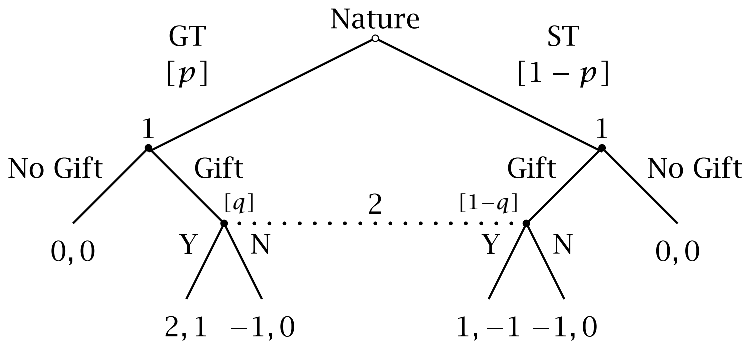


Figure 7: The Gift-Giving Game, II.

Player 2's expected payoff from Y is $q(1) + (1 - q)(-1) = 2q - 1$ and her expected payoff from N is $q(0) = (1 - q)(0) = 0$. Therefore, player 2 will accept whenever $2q - 1 > 0$, or $q > 1/2$, reject whenever $q < 1/2$, and indifferent between accepting and rejecting whenever $q = 1/2$. Her sequentially rational response now depends on her beliefs but we still don't know where these beliefs come from. How do we determine what q is going to be?

Requirements 1 and 2 ensure that players have beliefs and behave optimally given these beliefs. However, they do not require that these beliefs be reasonable. We now investigate how players can form these beliefs in a reasonable way.

2.3 Consistent Beliefs

In equilibrium, player 2's updated belief must be consistent with the probability distribution induced by any chance events and player 1's strategy. That is, we require that player 2 update her prior using information consistent with what she thinks player 1's strategy is. For example, in the gift-giving game, if player 2 thinks player 1 is playing the strategy NG, then if player 2 ever observes a gift offer, she must conclude that the book is the Star Trek manual, and therefore $q = 0$.

Recall that for a given equilibrium, an information set is **on the equilibrium path** if it is reached with positive probability if the game is played according to the equilibrium strategies, and is **off the equilibrium path** if it is certain that it cannot be reached.

Consider the $\langle NN, N\rangle$ equilibrium in the game from Fig. 5 (p. 5). If the game is played according to these strategies, then player 2's information set is never reached, and so it is off the equilibrium path. If, however, player 1 plays G with positive probability at least at one of his information sets (e.g. he plays a mixed strategy when the book is the game theory reference, in which he offers the gift with probability $1/2$), then the information set is on the equilibrium path because there is some positive probability that it will actually be reached if the game is played according to these strategies.

Along the equilibrium path, beliefs can be easily updated via Bayes rule from the strategies. That is, given player 2's prior belief and player 1's strategy, player 2 can observe player 1's action and form a posterior belief; i.e. a belief conditional on the observable action of player 1. The belief gives an intuitive answer to the question: Given that the information set is reached, what is the probability of being at a particular decision node?

Consider the game in Fig. 7 (p. 10). Player 2's prior belief about the book being GT is p. Suppose player 1 plays a mixed strategy, where he gives a gift with probability α_G if the book is GT, gives a gift with probability α_S if the book is ST. What must q be?

It is the probability of the book being GT given that the gift was offered. The probability that the gift is offered is the sum of the probabilities that it is offered depending on its type: $p\alpha_G + (1-p)\alpha_S$. The probability that it is offered if it is GT, is $p\alpha_G$ (that is, the probability that GT is selected by Nature times the probability that player 1 offers GT as a gift). Putting these things together gives us the posterior

$$q = \frac{p\alpha_G}{p\alpha_G + (1-p)\alpha_S},$$

which we got from Bayes rule.

Similarly, suppose there is a mixed strategy equilibrium in the game in Fig. 3 (p. 4), where player 1 chooses U with probability q, D with probability r, and D with probability $1-q-r$. In this case, p will be $p = q/(q+r)$. More generally, we have the following requirement.

Requirement 3 (Weak Consistency). Beliefs must be **weakly consistent** with strategies. That is, they are obtained from strategies and observed actions via Bayes rule whenever possible.

With the ideas of consistent beliefs and sequentially rational strategies, we have the following important result.

Theorem 2. *Suppose σ is an equilibrium in behavior strategies in an extensive form game with perfect recall. Let $h \in \mathcal{H}$ be an information set that occurs with positive probability under σ, and let π be a vector of beliefs that is weakly consistent with σ. Then σ is sequentially rational for player i at information set h with beliefs π.*

That is, at information sets that are visited with positive probability in an equilibrium, the equilibrium strategies are always sequentially rational. This means that if all information sets are reached with positive probability in some Nash equilibrium, then the equilibrium strategies must be sequentially rational. This implies that **in these cases, nothing is lost by the strategic form: The actions specified by the Nash equilibrium strategy profile for the strategic form representation would still be sequentially rational to carry out in the extensive form as well.** This should not be too surprising: the examples we gave had problems because the incredible actions occurred at information sets off the equilibrium path.

To illustrate the result of the theorem, recall Myerson's card game whose solution expressed in behavior strategies is:

$$\left\langle \left(\left({}^2\!/_3[F], {}^1\!/_3[R] \right), r \right), \left({}^2\!/_3[m], {}^1\!/_3[p] \right) \right\rangle .$$

All information sets are reached with positive probability in this strategy profile, and so the theorem tells us that these strategies must be sequentially rational. Let's verify that this is the case and check that each player would actually want to choose the actions specified by the strategies if the moves were not planned in advance.

Consider information set c for player 1 (chance has chosen the winning red card). If he raises, he would get at least 1 (if player 2 passes) and will get at most 2 (if player 2 meets). There is no gain from folding because it yields at most 1. Therefore, choosing r at this node is sequentially rational. Consider now his information set b (chance has chosen the losing black card). If player 1 raises, he would get $+1$ if player 2 passes, and -2 if player 2 meets. Under the assumption that player 2 implements her equilibrium strategy, the expected payoff will be $1/3 \times 1 + 2/3 \times -2 = -1$. If the player folds, his expected payoff is -1 also. Hence, player 1 is indifferent between raising and folding at this information set and must be willing to randomize. In other words, mixing at this information set is sequentially rational, just like the equilibrium strategy specifies.

Consider now player 2's information set h. Let q_b denote the left node (that is, the node that follows player 1 raising if the card is black), and let q_c denote the right node (that is, the node that follows player 1 raising if the card is red). We first calculate the probabilities of reaching these nodes under σ given Nature's moves:

$$\begin{aligned}
P(q_b|\sigma) &= (0.5) \times \sigma_1(R|b) = (0.5) \times 1/3 = 1/6, \\
P(q_c|\sigma) &= (0.5) \times \sigma_1(r|c) = (0.5) \times 1 = 0.5.
\end{aligned}$$

We now apply Bayes rule to derive the conditional probabilities; that is, the belief that the game has reached q_b given that player 2's information set was reached, denoted by $\pi_2(q_b|h)$; and the belief that the game has reached q_c given that player 2's information set was reached, denoted by $\pi_2(q_c|h)$:

$$\begin{aligned}
\pi_2(q_b|h) &= \frac{P(q_b|\sigma)}{P(q_b|\sigma) + P(q_c|\sigma)} = \frac{1/6}{1/6 + 0.5} = 0.25 \\
\pi_2(q_c|h) &= \frac{P(q_c|\sigma)}{P(q_b|\sigma) + P(q_c|\sigma)} = \frac{0.5}{1/6 + 0.5} = 0.75.
\end{aligned}$$

Of course, as required $\pi_2(q_b|h) + \pi_2(q_c|h) = 1$ because the belief must specify a valid probability distribution; that is, given that the information set was reached, it must be the case that exactly one node in it was reached. We conclude that given player 1's strategy and the chance moves by Nature, player 2's consistent belief given that her information set was reached must be that she is at node q_b with probability 0.25, and at node q_c with probability 0.75. With these consistent beliefs, we can now compute player 2's expected payoffs from her pure strategies. If she meets, she would get $\pi_2(q_b|h) \times 2 + \pi_2(q_c|h) \times -2 = 0.25 \times 2 + 0.75 \times -2 = -1$. If she passes, she would get -1 in either case. Hence, given her beliefs, player 2 is indifferent between meeting and passing because both strategies are sequentially rational. She can therefore mix in equilibrium, as prescribed by her equilibrium strategy.

2.4 Beliefs After Zero-Probability Events

You may have detected some hand-waving in Requirement 3 in the "whenever possible" clause. You would be right. How do we update beliefs in the strategy profile $\langle NN, N\rangle$ in Fig. 5 (p. 5)? The probability of reaching player 2's information set is 0 if these strategies are followed. We cannot use Bayes rule in such situations because it would involve division by 0. However, player 2's belief is still meaningful under these conditions: it is the belief when she is "surprised" by being offered a gift. This problem only arises off the equilibrium path, never on it (because along the equilibrium path there is never a zero probability of reaching an information set).

In the case when Bayes rule does not pin down posterior beliefs, *any* beliefs are admissible. This means that every action can be chosen as long as it is sequentially rational for some belief. Notice that in the gift game, N is not sequentially rational for any possible belief, and so it still would not be chosen. This is because N is strictly dominated by Y.

To help illustrate these ideas, consider the following motivational example in Fig. 8 (p. 13), where three players play a game that can end if player 1 opts out. Let p denote player 3's belief that he is at the left node in his information set conditional on this set being reached by the path of play.

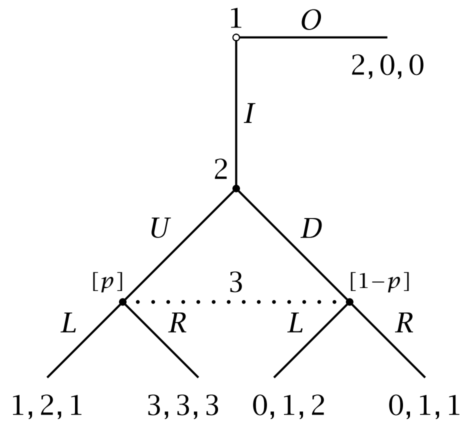


Figure 8: The Bigger Horsey Game.

Consider first the strategy profile (I, U, R). It is a Nash equilibrium of the game, and these strategies along with $p = 1$ satisfy Requirements 1 through 3 because there is no information set off the equilibrium path.

Consider now the strategy profile (O, U, L) and the belief $p = 0$. These strategies satisfy Requirements 1 and 2 but fail Requirement 3. The strategies do form a Nash equilibrium because no player wants to deviate unilaterally. Player 3 has a belief and acts optimally given that belief, and players 1 and 2 both act optimally given the strategies of the other players.

However, this Nash equilibrium fails the requirement of consistent beliefs. Player 3's belief is inconsistent with player 2's strategy. However, since the information set is never reached, Nash equilibrium cannot pin that down. Requirement 3, however, does because it forces player 3 to form beliefs consistent with the other players' strategies. Since player 2 chooses U in this profile, the only belief player 3 can hold is $p = 1$, in which case his strategy of playing L is no longer sequentially rational, and so it fails Requirement 2.

Let's now modify this example to demonstrate the "whenever possible" clause. Consider the game in Fig. 9 (p. 14), where player 2 has now the option to quit, ending the game as well.

Consider some equilibrium where player 1's optimal strategy was to play O, in which case player 3's information set is off the equilibrium path, as above. However, now Requirement 3 may not pin down player 3's beliefs because player 2 may choose Q, and so the probability of reaching player 3's information set conditional on this strategy is zero, and the weak consistency requirement does not restrict the beliefs there. We can assign anything we want there (a rather dissatisfying thing to do). Other refinements do put additional restrictions to handle these cases. Note, of course, that if player 2 chooses U with probability q_1, D with probability q_2, and Q with probability $1 - q_1 - q_2$, then Requirement 3 does have a bite because now player 3's information set is reached with positive probability given player 2's strategy, and so we require that $p = q_1/(q_1 + q_2)$ whenever $q_1 + q_2 > 0$.

What beliefs players have after zero-probability events is not a minor technical issue, but

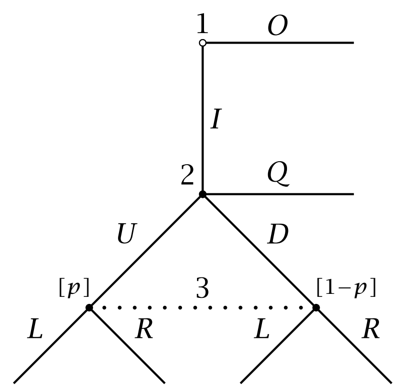


Figure 9: The Bigger Horsey Game, II.

an extremely important question, and much research in game theory has been directed at deciding what sort of beliefs are "reasonable" to have.

2.5 Perfect Bayesian Equilibrium

We have everything in place to define our solution concept, which is a stronger version of Nash equilibrium; i.e. it eliminates certain Nash equilibria that fail the additional requirements. We shall call the pair of a strategy profile and a belief vector, (σ, π), an **assessment**.

Definition 3. An assessment (σ^*, π^*) is a **perfect Bayesian equilibrium** (PBE) if the strategies specified by the profile σ^* are sequentially rational given beliefs π^*, and the beliefs π^* are weakly consistent with σ^*.

A PBE is a set of strategies and beliefs such that, at any stage in the game, strategies are optimal given the beliefs, and the beliefs are obtained from the equilibrium strategies and observed actions via Bayes rule whenever possible. Beliefs are elevated to the level of strategies here, and the equilibrium consists not just of a strategy for each player but also includes a belief for each player at each information set where the player has to move. Before we insisted that players choose reasonable strategies, we now also require that they hold reasonable beliefs.

The definition of PBE is circular in the sense that strategies must be optimal given beliefs and beliefs are derived from the strategies. This means that we must solve for strategies and beliefs simultaneously, like a system of equations. Sometimes this is quite involved, and we shall spend quite a bit of time practicing different ways of approaching these games. However, at least we do know that if we look for PBE, we shall find at least one in every game we are likely to solve in this class. The following result establishes this claim.

Theorem 3. *If (σ, π) is a perfect Bayesian equilibrium of an extensive-form game with perfect recall, then σ is a mixed-strategy Nash equilibrium. For any finite extensive-form game, a perfect Bayesian equilibrium exists.*

This theorem tells us that the set of perfect Bayesian equilibria is really a subset of the set of Nash equilibria. That is, any PBE is also a Nash equilibrium. The converse, of course, is not true: there are Nash equilibria that are not PBE. However, the theorem guarantees that the additional restrictions will not eliminate all Nash equilibria from consideration. That is, each finite game will have at least one PBE. This, as you can imagine, is very important if we want to solve games.

So remember, a PBE is a Nash equilibrium where strategies are sequentially rational given the beliefs, and the beliefs are weakly consistent with these strategies (updated via Bayes rule whenever possible).

Going back to our Gift-Giving Game in Fig. 7 (p. 10), recall that player 2's sequentially rational strategy is to accept whenever $q > 1/2$, reject whenever $q < 1/2$, and either one (including mixtures) otherwise. Let player 1's strategy be denoted by (rs), where r is the probability of offering the Game Theory book, and s is the probability of offering the Star Trek manual. Bayes rule then yields player 2's posterior belief:

$$q = \frac{pr}{pr + (1-p)s}.$$

To find the PBE, we must find mixing probabilities for the two players that are sequentially rational and that are also such that q is consistent with player 1's equilibrium strategy. Suppose first that player 1's strategy is completely mixed; that is, he randomizes at both information sets, so $r, s \in (0,1)$. Take an arbitrary (possibly degenerate) mixed strategy for player 2 in which she accepts with probability $\alpha \in [0,1]$. Since player 1 is willing to mix, he must be indifferent between offering the gift and not offering it at both information sets. Not offering gives him a payoff of 0 in either case. Offering, on the other hand, yields a payoff $U_1(G|\text{GT}) = 2\alpha + (1-\alpha)(-1)$ if the book is on Game Theory and $U_1(G|\text{ST}) = \alpha + (1-\alpha)(-1)$ if it is on Star Trek. Observe now that because player 1 must be indifferent to mix, it follows that $U_1(G|\text{ST}) = U_1(NG|\text{ST})$, or

$$\alpha + (1-\alpha)(-1) = 0 \Rightarrow 2\alpha - 1 = 0 \Rightarrow \alpha = \frac{1}{2}.$$

That is, if player 1 is indifferent between offering the Star Trek manual and not offering it, it must be the case that player 2 will accept the offer with probability exactly equal to $1/2$. This now implies that

$$U_1(G|\text{GT}) = 2(1/2) + (1 - 1/2)(-1) = 1 - 1/2 = 1/2 > 0 = U_1(NG|\text{GT}).$$

In other words, $\alpha = 1/2$, which must hold for player 1 to mix in if he holds the Star Trek manual, also ensures that he cannot possibly mix if he holds the Game Theory book. This contradicts the supposition that player 1 mixes in both cases. We conclude that if player 1 mixes on the Star Trek manual in equilibrium, he must be offering the Game Theory book for sure.

Since we now know that $s^* > 0 \Rightarrow r^* = 1$, let's see if there is a PBE with these properties. From the discussion above, we know that this equilibrium requires $\alpha^* = 1/2$ or else player 1 would not randomize with the Star Trek manual. This now implies that $q = 1/2$ or else player 2 would not randomize in her acceptance decision. Putting everything together yields:

$$q = \frac{1}{2} = \frac{p(1)}{p(1) + (1-p)s^*} \Rightarrow s^* = \frac{p}{1-p}.$$

We need to ensure that s^* is a valid mixing probability; that is, we must make sure that $s^* \in (0,1)$. It is clearly positive because $p \in (0,1)$. To ensure that $s^* < 1$, we also need $p < 1/2$. Hence, we found a PBE. Writing it in behavioral strategies yields the following:

$$\left\langle \left(r^* = 1, s^* = \frac{p}{1-p} \right), \alpha = 1/2 \right\rangle \text{ provided } p < \frac{1}{2}.$$

This equilibrium is intuitive: since player 1 always offers the Game Theory book and only sometimes offers the Star Trek manual, player 2 is willing to risk accepting his offer. Of course, her estimate about this risk depends on her prior belief. Player 1's strategy is precisely calibrated to take into account this belief when he tries to bluff player 2 into accepting what he knows is a gift she would not want. Note that $p < 1/2$ is a necessary condition for this equilibrium. As we shall see shortly, if player 2's priors are too optimistic, she would accept the offer for sure, in which case (we would expect) player 1 to offer her even the Star Trek manual for sure.

Observe now that player 2 has learned something from player 1's strategy that she did not know before: her equilibrium posterior belief is $q = 1/2 > p$. That is, she started out with a prior which assigned less than 50% chance to the gift being a Game Theory book and then updated this belief to 50% upon seeing the gift being offered. However, she still is not sure just what type of gift she is being offered. Player 1's strategy is called *semi-separating* because player's action allow player 2 to learn something, but not everything, about the information he has: she can "separate" the Game Theory gift from the Star Trek manual only partially.[3] The residual uncertainty player 2 has is a common feature of the other player choosing a semi-separating strategy.

We have exhausted the possibilities in which player 1 mixes when he has the Star Trek manual. Only two possible types of strategies remain: he either always offers it or never does. Suppose first that player 1 always offers the Star Trek manual in equilibrium, so $s^* = 1$. From our calculations above, we know that this means player 2 would have to accept his offer with $\alpha \geq 1/2$, which in turn implies that $q \geq 1/2$ as well. If player 2 accepts with probability at least as high as $1/2$, then player 1 will always offer the Game Theory book: $U_1(G|\text{GT}) = 3\alpha - 1 > 0$ for any $\alpha > 1/3$. This now means that player 1 always offers the gift regardless of its type, which implies $q = p$. Since we require that $q \geq 1/2$, it follows that this equilibrium only exists if $p \geq 1/2$. Hence, we found a PBE for that range of priors:

$$\langle (r^* = 1, s^* = 1), \alpha = 1 \rangle \text{ provided } p \geq \frac{1}{2}.$$

Note that if $p = 1/2$, then player 2 is indifferent between accepting and rejecting, so she can mix with any probability as long as $\alpha \geq 1/2$, so there is a continuum of PBE in this case. However, requiring that the prior equal a particular value is an extremely demanding condition and these solutions are extremely fragile: the smallest deviation from $p = 1/2$ would immediately produce one of the PBE we identified above. Normally, we would ignore solutions that depend on knife-edge conditions like that. It is important to note that whereas $p = 1/2$ is a knife-edge condition we can ignore, $q = 1/2$ in our semi-separating PBE is not. Unlike the prior, the posterior probability is strategically induced by the behavior of the player.

In this equilibrium, player 1 is playing a *pooling* strategy because he "pools" on the same action (offering the gift) no matter what he knows about the gift's type. Not surprisingly, whenever a player uses a pooling strategy, his opponent cannot learn anything from the behavior she observes. As we have seen, her posterior is exactly equal to her prior.

Suppose now that player 1 never offers the Star Trek manual, so $s^* = 0$, but offers the Game Theory book with positive probability, so $r^* > 0$. In this case, $q = 1$, and player 2's sequentially rational response is to accept for sure. However, if she accepts the offer with certainty, then player 1 would prefer to offer even the Star Trek manual: doing so would net

[3] For this reason, some authors call the strategies "partially separating" or "partially pooling."

him a payoff of 1 while not offering saddles him with 0. In other words, there can be no equilibrium of this type.

In this case, the player 1's strategy is *separating* because it allows player 2 to infer what he knows with certainty. Not surprisingly, if he allowed her to do that, he would be at a disadvantage. It is not difficult to see that separating in a way that convinces player 2 that the offer contains the Game Theory book cannot be optimal: she would accept the offer given this belief but then player 1 would have an incentive to offer the Star Trek manual as well, which means player 2 cannot sustain her belief that she is offered the Game Theory book for sure.

Conversely, player 1 could induce a posterior belief that the book is the Star Trek manual for sure by playing the separating strategy $(r^* = 0, s^* = 1)$; that is, he never offers the Game Theory book and always offers the Star Trek manual. But in this case, player 2 will always reject the offer, in which case player 1 will be strictly better off not offering the Star Trek manual. So this cannot be an equilibrium either. This game has no separating equilibria.

The final scenario to examine is $r^* = s^* = 0$. This leaves player 2's information set off the equilibrium path of play, so Bayes rule cannot pin down her posterior beliefs. Can we sustain these strategies in equilibrium? To induce player 1 never to offer even the Game Theory book, it must be the case that $U_1(G|\text{GT}) \le U_1(NT|\text{GT}) \Rightarrow 3\alpha - 1 \le 0 \Rightarrow \alpha \le {}^1/_3$. That is, player 2 must threaten that she will reject the offer with at least ${}^2/_3$ probability. This threat will be credible if she believed the gift was bad with probability at least ${}^1/_2$. In other words, if $q \le {}^1/_2$, then player 2 can credibly reject the offer. Observe now that this is sufficient to deter player 1 from offering the Star Trek manual as well: $U_1(G|\text{ST}) = 2\alpha - 1 < 2({}^1/_3) - 1 = -{}^1/_3 < 0 = U_1(NG|\text{ST})$. Therefore, we have multiple PBE that all take the same form:

$$\langle (r^* = 0, s^* = 0), \alpha = 0 \rangle \text{ with } q \le {}^1/_2.$$

We have some latitude in specifying off-the-equilibrium-path beliefs here but as long as $q \le {}^1/_2$, these strategies can be sustained in PBE. Note that technically when $q = {}^1/_2$, any $\alpha \le {}^1/_3$ would do the trick. However, as a knife-edge case that does not change the equilibrium path of play, we would normally ignore this.

Observe also that this PBE exists regardless of the prior. In this situation, we have *multiple equilibria*. It is important to realize that the semi-separating and the other pooling equilibrium in which player 1 always makes an offer are not "multiple equilibria" in that sense: depending on the value of the exogenous parameter p, the equilibrium takes either the partially separating or the pooling form. If it weren't for the pooling PBE where player 1 never offers the gift, this game would have a unique solution for any value of this exogenous parameter. Unfortunately, with this second pooling PBE, this is no longer the case: the game now has at least two solutions (in terms of the equilibrium path of play) over the entire range of p, an indeterminacy that we would have to address somehow.

3 Backward Induction

Before learning how to characterize PBE in arbitrary games of imperfect information, we study how to use backward induction to compute PBE in games of complete information. In games of complete and perfect information, we can apply this technique to find the "rollback" equilibrium, which we then generalize to "subgame-perfect" equilibrium. All of these are PBE but do not require the entire belief machinery.

3.1 Rollback Equilibrium

Computing PBE in games of complete and perfect information is very easy and one does not need to introduce the machinery of beliefs at all because all information sets are singletons. Sequential rationality will suffice (recall that the conditional beliefs would always require that the probability of being at a node in a singleton information set is exactly 1).

These games can be solved by **backward induction**, a technique which involves starting from the last stage of the game, determining the last mover's best action at his information set there, and then replacing the information set with the payoffs from the outcome that the optimal action would produce. Continuing in this way, we work upwards through the tree until we reach the first mover's choice at the initial node.

In 1913 Zermelo proved that chess has an optimal solution. He reasoned as follows. Since chess is a finite game (it has quite a few moves, but they are not infinite), this means that it has a set of penultimate nodes. That is, nodes whose immediate successors are terminal nodes. The optimal strategy specifies that the player who can move at each of these nodes chooses the move that yields him the highest payoff (in case of a tie he makes an arbitrary selection). Now, the optimal strategies specify that the player who moves at the nodes whose immediate successors are the penultimate nodes chooses the action which maximizes his payoff over the feasible successors given that the other player moves there in the way we just specified. We continue doing so until we reach the beginning of the tree. When we are done, we will have specified an optimal strategy for each player.

These strategies constitute a Nash equilibrium because each player's strategy is optimal given the other player's strategy. In fact, these strategies also meet the stronger requirements of subgame perfection, which we shall examine in the next section. (Kuhn's paper provides a proof that any finite extensive form game has an equilibrium in pure strategies. It was also in this paper that he distinguished between mixed and behavior strategies for extensive form games.) Hence the following result:

THEOREM 4 (Zermelo 1913; Kuhn 1953). *A finite game of perfect information has a pure strategy Nash equilibrium.*

We can find this equilibrium by backward induction in the following way. Start with the penultimate node, and determine the sequentially rational strategy for the player who gets to move there. Replace the node with the outcome produced when that player chooses this strategy. Continue working backward until you reach the initial node. The sequentially rational strategies you now have constitute the PBE of this game, which is called a **rollback equilibrium**.

Since Selten's game in Fig. 2 (p. 3) is one of complete and perfect information, we can apply backward induction to find its rollback equilibrium. At her information set, player 2 would choose L. This reduces player 1's choices between D (which, given player 2's strategy would yield 3) and U, which yields 2. Therefore, player 1 would choose D. The rollback equilibrium of this game is $\langle D, L\rangle$. As before, this is also Nash and PBE.

Consider now the game in Fig. 10 (p. 19).

What are the Nash equilibria of this game? As usual, convert this to strategic form, as shown in Fig. 11 (p. 19). (We shall keep the non-reduced version to illustrate a point.) The Nash equilibria in pure strategies are $\langle(\sim e, a), r\rangle$, $\langle(\sim e, \sim a), r\rangle$, and $\langle(e, a), \sim r\rangle$.

Applying backward induction produces the rollback equilibrium $\langle(\sim e, \sim a), r\rangle$ illustrated in Fig. 12 (p. 19). This eliminates two of the pure-strategy Nash equilibria, and demonstrates

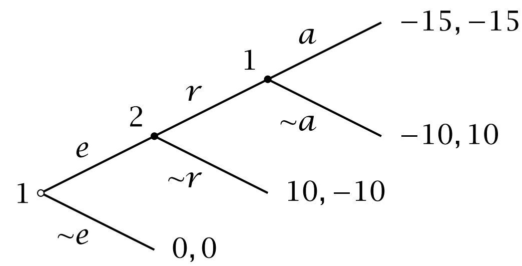


Figure 10: Basic Escalation Game.

		Player 2: r	Player 2: $\sim r$
Player 1	(e, a)	$-15, -15$	$10, -10$
	$(e, \sim a)$	$-10, 10$	$10, -10$
	$(\sim e, a)$	$0, 0$	$0, 0$
	$(\sim e, \sim a)$	$0, 0$	$0, 0$

Figure 11: The Strategic Form of the Escalation Game.

why it is extremely important that strategies specify moves even at information sets that would not be reached if the strategy is followed.

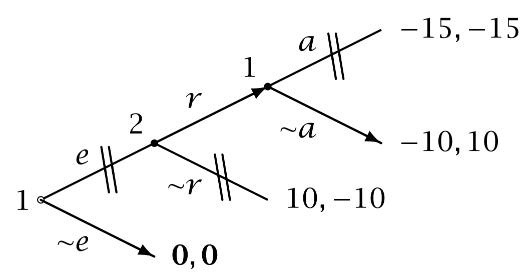


Figure 12: The Rollback Equilibrium of the Escalation Game.

The only reason why $\sim e$ is sequentially rational at player 1's first information set is because player 2's sequentially rational strategy prescribes r, which in turn is only rational because she expects player 1 to choose $\sim a$ at his last information set, where this is the sequentially rational choice. In other words, the optimality of player 1's initial action depends on the optimality of his action at his second information set. This is precisely why we cannot determine optimality of strategies unless they specify what to do for all information sets. Note that in this case, the second information set is not reached if the strategy $(\sim e, \sim a)$ is followed, but we still need to know the action there.

The problem with the Nash equilibrium profiles $\langle(\sim e, a), r\rangle$ and $\langle(e, a), \sim r\rangle$ is that they leave information sets off the equilibrium path of play and so Nash optimality cannot pin down behavior at sets that are never reached. For example, $\langle(e, a), \sim r\rangle$ leaves player 1's second information set off the path of play, which causes the strategy to miss the fact that a is not sequentially rational at that set. This incredible threat rationalizes player 2's choice of $\sim r$ causing her to take an action that leaves this information set off the path. Similarly, $\langle(\sim e, a), r\rangle$ leaves both player 1's second information set and player 2's information set off the path of play and causes the strategies to miss two problems: player 1's choice of a is not sequentially rational at his second information set, and given that choice, player 2's choice of r is not rational either! In other words, since the equilibrium path of play does not reach some continuation games, Nash cannot pin down optimal behavior there.

Let's apply backward induction to the game in Fig. 13 (p. 20). There are three penultimate information sets, all of them for player 2. For each of these sets we determine player 2's sequentially rational strategy. Because all information sets are singletons, we do not need to specify conditional beliefs (they are trivial and assign probability 1 to the single node in the information set). So, if the $(2,0)$ set is ever reached, then player 2 is indifferent between accepting and rejecting. Any of these actions is plausible, so player 2's sequentially rational strategy at $(2,0)$ includes both actions. In other words, player 2 will be willing to mix at this information set. On the other hand, player 2's sequentially rational strategy at $(1,1)$ is unique: it is always better to accept. And so in the case of $(0,2)$. Therefore, player 2 has only two credible pure strategies, yyy and nyy.

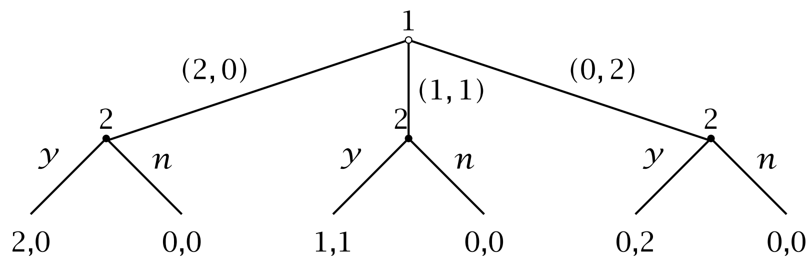


Figure 13: The Two-Way Division Game, Perfect Information.

Given player 2's best responses to each of his actions, player 1 can now optimize his own behavior. He knows that if he chooses $(2,0)$, then player 2 can either accept of reject, if he offers $(1,1)$ or $(0,2)$, then player 2 will always accept. Given these conditions, player 1 will never offer $(0,2)$ in equilibrium because it yields him a payoff of 0. Given these strategies, player 1 has two optimal actions, either $(2,0)$ if he thinks player 2 will choose the first strategy and $(1,1)$ if he thinks that player 2 will choose the second. Thus, we conclude that the game has only two rollback equilibria in pure strategies: $\langle(2,0), yyy\rangle$ and $\langle(1,1), nyy\rangle$.[4]

Kuhn's theorem makes no claims about uniqueness of the equilibrium. Indeed, as we saw in the example above, the game has two rollback equilibria. However, it should be clear that if no player is indifferent between any two outcomes, then the equilibrium will be unique. Note that all rollback equilibria are Nash equilibria. (The converse, of course, is not true. We have just gone to great lengths to eliminate Nash equilibria that seem implausible.)

3.2 Subgame-Perfect Equilibrium

If you accept the logic of backward induction, then the following discussion should seem a natural extension. Consider the game in Fig. 14 (p. 21). Here, neither of player 2's choices is

[4]Player 2 is indifferent between accepting and rejecting after $(2,0)$ offer, so she can mix after receiving it. If she accepts with probability greater than $1/2$, then player 1 is better off demanding $(2,0)$, and we have a continuum of SPE there. If, however, she accepts with probability less than $1/2$, then player 1 is better off demanding $(1,1)$ instead (for another continuum of SPE). If she accepts exactly with probability $1/2$, then player 1 is indifferent between $(2,0)$ and $(1,1)$, so he can mix between these two strategies as well (for another continuum of SPE).

dominated at her second information set: she is better off choosing D if player 1 plays A and is better off choosing C if player 1 plays B. Hence, we cannot apply rollback.

However, we can reason in the following way. The game that begins with player 1's second information set—the one following the history (D, R)—is a zero-sum simultaneous move game. We have seen similar games, e.g., MATCHING PENNIES. The expected payoffs from the unique mixed strategy Nash equilibrium of this game are $(0, 0)$. Therefore, player 2 should only choose R if she believes that she will be able to outguess player 1 in the simultaneous-move game. In particular, the probability of obtaining 2 should be high enough (in outweighing the probability of obtaining -2) that the expected payoff from R is larger than 1 (the payoff he would get if he played L). This can only happen if player 2 believes she can outguess player 1 with a probability of at least $3/4$, in which case the expected payoff from R will be at least $3/4(2) + 1/4(-2) = 1$. But, since player 2 knows that player 1 is rational (and therefore just as cunning as she is), it is unreasonable for her to assume that she can outguess player 1 with such high probability. Therefore, player 2 should choose L, and so player 1 should go D. The equilibrium obtain by backward induction in the game in Fig. 14 (p. 21), then, is $\langle (D, 1/2[A]), (L, 1/2[C]) \rangle$.

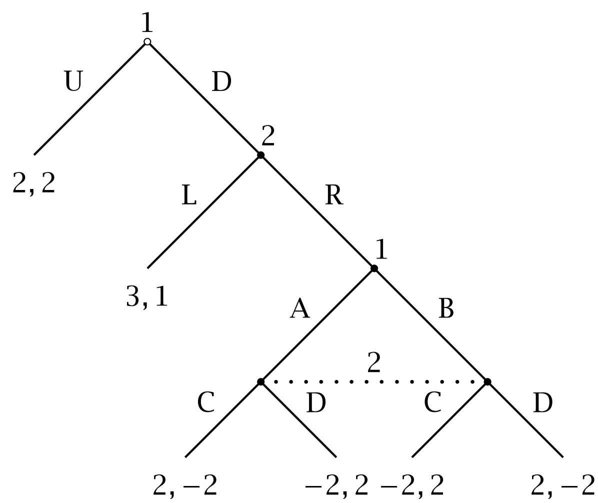


Figure 14: The Fudenberg & Tirole Game.

This now is the logic of subgame perfection: replace every "proper subgame" of the tree with one of its Nash equilibrium payoffs and perform backward induction on the reduced tree.[5] For the game in Fig. 14 (p. 21), once we replace the subgame that starts at player 1's second information set with the Nash equilibrium outcome, the game becomes the one in Fig. 2 (p. 3), which we have already analyzed and for which we found that the backward-induction equilibrium is $\langle D, L \rangle$.

We were a little vague in the preceding paragraph. Before we formally define what a subgame perfect equilibrium is, we must define what constitutes a "proper subgame." It really isn't hard: a proper subgame is any part of a game that can be analyzed as a game itself.

DEFINITION 4. A **proper subgame** $\boldsymbol{G}$ of an extensive-form game Γ consists of a *single* decision node and all its successors in Γ with the property that if $x' \in G$ and $x'' \in h(x')$, then $x'' \in G$ as well. The payoffs are inherited from the original game.

[5]If the game has multiple Nash equilibria, then players must agree on which of them would occur. We shall examine this weakness in the following section.

That is, x' and x'' are in the same information set in the subgame if and only if they are in the same information set in the original game. The payoffs in the subgame are the same as the payoffs in the original game only restricted to the terminal nodes of the subgame. Note that the word "proper" does not mean strict inclusion as in the term "proper subset." Any game is always a proper subgame of itself.[6]

Proper subgames are quite easy to identify in a broad class of extensive form games. For example, in games of complete and perfect information, every information set (a singleton) begins a proper subgame (which then extends all the way to the end of the tree of the original game). Each of these subgames represents a situation that can occur in the original game.

On the other hand, splitting information sets in games of imperfect information produces subgames that are not proper because they represent situations that cannot occur in the original game. Consider, for example, the game Fig. 15 (p. 22) and two candidate subgames.

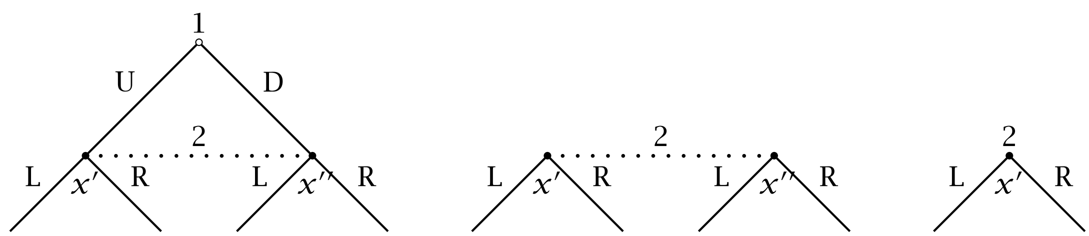


Figure 15: A Game with Two "Improper" Subgames.

The two subgames to the right of the original game are not proper. The first one fails the requirement that a proper subgame begin with a single decision node. The second one fails the requirement that if two decision nodes are in the same information set in the original game, they must also be in the same information set in the proper subgame.

The reasons for these restrictions are intuitive. In the first case, player 2 needs to know the relative probabilities for the decision nodes x' and x'' but the "game" specification does not provide these probabilities. Therefore, we cannot analyze this situation as a separate game. In the second case, player 2 knows that player 1 did not play D, and so has more information than in the original game, where he did not know that.

To make things a little easier, here are some guidelines for identifying subgames. A subgame (a) always starts with a single decision node, (b) contains all successors to that node, and (c) if it contains a node in an information set, then it contains all nodes in that information set. (Never split information sets.)

Now, given these restrictions, the payoffs conditional on reaching a proper subgame are well defined. We can therefore test whether strategies are a Nash equilibrium in the proper subgame as we normally do. This allows us to state the new solution concept.

Definition 5. A behavior strategy profile σ of an extensive form game is a **subgame perfect equilibrium** if the restriction of σ to G is a Nash equilibrium for every proper subgame G.

You should now see why it was necessary to define the behavior strategies: some proper subgames (e.g. the one in F&T's Game from Fig. 14 (p. 21)) have subgames where the Nash

[6]Rasmusen departs from the convention in his book *Games and Information*, where he defines a proper subgame to mean strict inclusion, and so he excludes the entire game from the set. We shall follow the convention.

equilibrium is in mixed strategies, which requires that players be able to mix at each information set. (You should at this point go over the difference between mixed and behavior strategies in extensive-form games.)

This now allows us to solve games like the one in Fig. 14 (p. 21). There are three proper subgames: the entire game, the subgame beginning with player 2's information set, and the subgame that includes the simultaneous moves game. We shall work, as we did with backward induction, our way up the tree. The smallest proper subgame has a unique Nash equilibrium in mixed strategies, where each player chooses one of the two available actions with the same probability of .5. Given these strategies, each player's expected payoff from the subgame is 0. This now means that player 2 will choose L at her information set because doing so nets her a payoffs strictly larger than choosing R and receiving the expected payoff of the simultaneous-moves subgame. Given that player 2 chooses L at her information set, player 1's optimal course of action is to go D at the initial node. So, the subgame perfect equilibrium of this game is $\langle (D, 1/2), (L, 1/2) \rangle$.

Let's compare this to the normal and reduced normal forms of this extensive-form game; both of which are shown in Fig. 16 (p. 23).

Player 2

Player 1	LC	LD	RC	RD
UA	2, 2	2, 2	2, 2	2, 2
UB	2, 2	2, 2	2, 2	2, 2
DA	3, 1	3, 1	2, −2	−2, 2
DB	3, 1	3, 1	−2, 2	2, −2

	L	RC	RD
U	2, 2	2, 2	2, 2
DA	3, 1	2, −2	−2, 2
DB	3, 1	−2, 2	2, −2

Figure 16: The Normal and Reduced Normal Forms of the Game from Fig. 14 (p. 21).

The normal form game on the left has 4 pure strategy Nash equilibria: $\langle UA, RC \rangle$, $\langle UA, RD \rangle$, $\langle UB, RC \rangle$, and $\langle UB, RD \rangle$. The reduced normal form game has only two: $\langle U, RC \rangle$ and $\langle U, RD \rangle$. None of these are subgame perfect. However, the reduced form also has a Nash equilibrium in mixed strategies, $\langle \sigma_1^*, \sigma_2^* \rangle$, in which $\sigma_1^*(DA) = \sigma_1^*(DB) = 1/2$, and $\sigma_1^*(U) = 0$; while $\sigma_2^*(L) = 1$, and $\sigma_2^*(RC) = \sigma_2(RD) = 0$. The Nash equilibrium is

$$\langle (0, 1/2, 1/2), (1, 0, 0) \rangle$$

which is precisely the subgame-perfect equilibrium we already found.

At this point you should make sure you can find this mixed strategy Nash equilibrium. Suppose player 2 chooses RC for sure, then DB is strictly dominated, so player 1 will not use it. However, this now means RD strictly dominates RC for player 2, a contradiction. Suppose now that she chooses RD for sure. Then DA is strictly dominated, so player 1 will not use it. But now RC strictly dominates RD, a contradiction. Therefore, there is no equilibrium in which player 2 chooses either RC or RD for sure. Suppose player 2 puts positive weight on L and RC only. Then, DA is strictly dominant for player 1. However, 2's best response to DA is RD, a contradiction with supposition that she does not play it. Hence, no MSNE in which she plays only L and RC. Suppose now that she plays L and RD only. Then DB is strictly dominant, but player 2's best response to this is RC, a contradiction. Hence, no MSNE in which she plays only L and RD either. Suppose next that she plays RC and RD only. Then U is strictly dominant, and since player 2's payoff is the same against U regardless of the strategy she uses, we have a continuum of MSNE: $\langle U, \sigma_2(RC) \in (0,1), \sigma_2(RD) = 1 - \sigma_2(RC) \rangle$.

Suppose next she plays L for sure. Then player 1 is indifferent between DA and DB, each of which strictly dominates U, so he can mix with $\sigma_1(DA) \in (0,1)$ and $\sigma_1(DB) = 1 - \sigma_1(DA)$. Since player 2 must not be willing to use any of her other pure strategies, it follows that $U_2(\sigma_1(DA), RC) \leq 1 \Leftrightarrow \sigma_1(DA) \geq 1/4$, and $U_2(\sigma_1(DA), RD) \leq 1 \Leftrightarrow \sigma_1(DA) \leq 3/4$. Therefore, $\sigma_1(DA) \in [1/4, 3/4]$ are all admissible mixtures, and we have a continuum of MSNE. The subgame-perfect MSNE is among these: the one with $\sigma_1(DA) = 1/2$.[7]

As you can see, we found a lot of MSNE but only one of them is subgame-perfect. This reiterates the point that all SPE are Nash, while not all Nash equilibria are subgame-perfect. Note the different way of specifying the equilibrium in the extensive form and in the reduced normal form.

We can now state a very important result that guarantees that we can find subgame perfect equilibria for a great many games.

THEOREM 5. *Every finite extensive game with perfect information has a subgame perfect Nash equilibrium.*

To prove this theorem, simply apply backward induction to define the optimal strategies for each subgame in the game. The resulting strategy profile is subgame perfect.

Let's revisit our basic escalation game from Fig. 10 (p. 19). It has three subgames, shown in Fig. 17 (p. 24) and labeled I, II, and III. What are the pure-strategy Nash equilibria in all these subgames? We have already found the three equilibria of subgame I: $\langle(\sim e, a), r\rangle$, $\langle(\sim e, \sim a), r\rangle$, and $\langle(e, a), \sim r\rangle$. The Nash equilibrium of subgame III is trivial: $\sim a$. You should verify (e.g. by writing the normal form) that the Nash equilibria of subgame II are: $\langle a, \sim r\rangle$ and $\langle \sim a, r\rangle$.

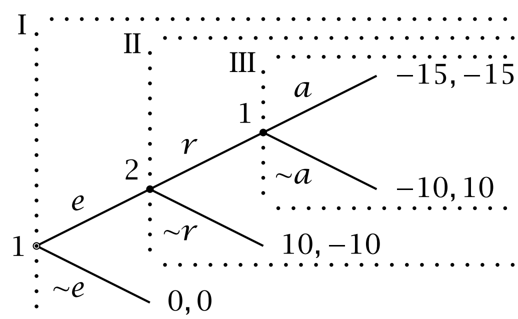


Figure 17: The Subgames of the Basic Escalation Game.

Of the three equilibria in subgame I (the original game), which ones are subgame perfect? That is, in which of these do the strategies constitute Nash equilibria in all subgames? The restriction of the strategies to subgame II shows that no strategy profile that involves anything other than the combinations $\langle a, \sim r\rangle$ and $\langle \sim a, r\rangle$ would be subgame perfect. This eliminates the Nash equilibrium profile $\langle(\sim e, a), r\rangle$ of the original game. Further, the restriction of the strategies to game III demonstrates that no profile that involves player 1 choosing anything other than $\sim a$ would be subgame perfect either. This eliminates the Nash equilibrium profile $\langle(e, a), \sim r\rangle$ of the original game. There are no more subgames to check, and therefore all remaining Nash equilibria are subgame perfect. There is only one remaining Nash equilibrium:

[7] I leave the final possibility as an exercise: what happens if player 2 puts positive weight on all three of her strategies?

$\langle(\sim e, \sim a), r\rangle$ and this is the unique subgame perfect equilibrium. Of course, it is the one we got from our backward induction method as well.

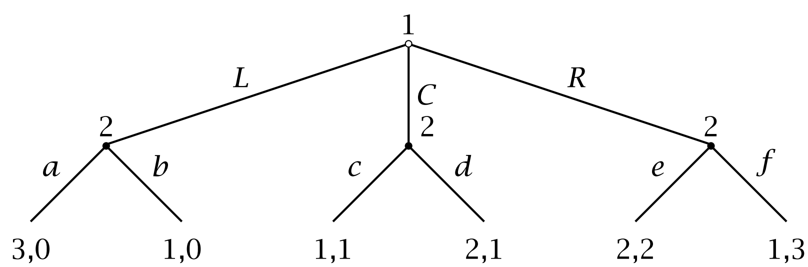
Subgame perfection (and backward induction) eliminates equilibria based upon non-credible threats and/or promises. This is accomplished by requiring that players are rational at every point in the game where they must take action. That is, their strategies must be optimal at every information set, which is is a much stronger requirement than the one for Nash equilibrium, which only demands rationality at the first information set. This property is, of course, **sequential rationality**, just as we defined it above.

4 Practice with Subgame Perfection

Note that nowhere in our definition of extensive form games did we restrict either the number of actions available to players at their decision nodes nor the number of decision nodes. For example, a player may have a continuum of actions or some terminal history may be infinitely long. If a player has a continuum of action at any decision node, then there is an infinite number of terminal histories as well. We must distinguish between games that exhibit some finiteness from those that are infinite.

If the length of the longest terminal history is finite, then the game has **finite horizon**. If the game has finite horizon and finitely many terminal histories, then the game is **finite**. Backward induction only works for games that are finite. Subgame perfection works fine for infinite games.

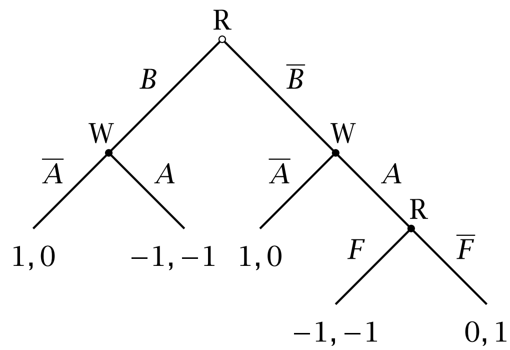
Let's begin by finding all subgame perfect equilibria in the following game.



Consider player 2's information set after history R. Playing f is strictly better than e, so any equilibrium strategy for player 2 will involve choosing f at that information set. However, in the information sets following histories L and C, player 2 is indifferent between her two actions, so she can pick either one from each pair. This yields four strategies for player 2: (acf), (adf), (bcf), and (bdf). Player 1's best response to any strategy for player 2 that prescribes a as the first action is to player L. Therefore, $\langle L, acf\rangle$ and $\langle L, adf\rangle$ are subgame perfect equilibria. Also, if player 2 is choosing b and d at the first two of her information sets, then player 1's best response is to player C. Therefore, $\langle C, bdf\rangle$ is another SPE. If player 2's strategy is (bcf), then player 1 is indifferent between L, C, and R. Therefore, $\langle L, bcf\rangle$, $\langle C, bcf\rangle$, and $\langle R, bcf\rangle$ are SPE. The game thus has six SPE in pure strategies. It has more in mixed strategies.

4.1 Burning a Bridge

The Red Army, having retreated to Stalingrad is facing the advancing *Wehrmacht* troops. If the Red Army stays, then it must decide whether to defend Stalingrad in the case of an attack or retreat across Volga using the single available bridge for the purpose. Each army prefers to occupy Stalingrad but fighting is the worst outcome for both. However, before the enemy can attack, the Red Army can choose to blow up the bridge (at no cost to it), cutting off its own retreat. Model this as an extensive form game and find all SPE.



Since this is a finite game of complete and perfect information, let's solve it by backward induction. Starting with the longest terminal history, $(\overline{B}, A)$, Red Army's optimal action is to retreat across the bridge, or $\overline{F}$. Given this action, the *Wehrmacht* strategy following history $(\overline{B})$ would be to attack, or A. Its strategy after history (B) is not to attack, or $\overline{A}$. Thus, the Germans' optimal strategy is $(\overline{A}, A)$. Given this strategy, the Red Army strictly prefers to burn the bridge. So, the unique SPE is $\left\langle (B, \overline{F}), (\overline{A}, A) \right\rangle$. The outcome is that the Red Army burns the bridge and the Germans don't attack.

This is an easy example that demonstrates a rather profound result of strategic interaction: if you limit your choices and do so in a way that is observable by the opponent, then you may obtain better outcomes. This is because unless the Red Army burns the bridge, it cannot credibly commit to fighting in order to induce the Germans not to attack. (Their threat to fight if attacked is not credible, and so deterrence fails.) However, by burning the bridge, they leave themselves no choice but fight if attack, even though they don't like it. This makes the threat to fight credible, and so the Germans are deterred.

You will see credible commitment problems in a wide variety of contexts in political science. Generally, a commitment is not credible if it will not be in the interest of the committing party to carry out its promises should it have to do so. (In our language, its threats/promises are not subgame perfect.) Limiting one's choices in an observable way may help in such situations.[8]

4.2 The Dollar Auction

Let's now play the following game. I have \$1 that I want to auction off using the following procedure. Each student can bid at any time, in 10 cent increments. When no one wants to bid further, the auction ends and the dollar goes to the highest bidder. Both the highest bidder and the second highest bidder pay their bids to me. Each of you has \$3.00 to bid with and you cannot bid more than that.

[What happened?]

[8]This is probably obvious by now but sometimes people get it completely wrong. Take the Trojans who tried to burn the Greeks' ships! Had they succeeded in doing so, this would have only made the Greeks fight so much harder. (They failed and so the Greeks sailed in apparent defeat, enabling them to carry out the wooden horse ruse.) William the Conqueror and Cortez got it right when the formed burned and the latter scuttled *their own* ships, forcing the soldiers to fight to the end and compelling some of the opposition to surrender.

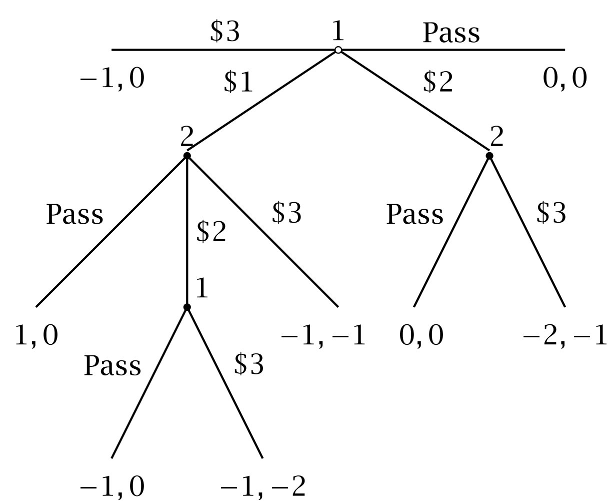
Let's analyze this situation by applying backward induction. We shall assume that it is not worth spending a dollar in order to win a dollar. Suppose there are only two players, 1 and 2. If 2 ever bids \$3.00 she will win the auction and lose \$2.00. If she bids \$2.90, then 1 has to bid \$3.00 in order to win. If 1's previous bid was \$2.00 or less, then 1 will not outbid 2's \$2.90 because doing so (and winning) means losing \$2.00 instead of whatever 1's last bid is, which is at most that (this is where the assumption comes into action). But the same logic applies for 2's bid of \$2.80 because 1 cannot bid \$2.90 and expect to win: in this case 2 will do strictly better by outbidding him with \$3.00 and losing only \$2.00 instead of \$2.80. Therefore, 1 can only win by bidding \$3.00 but this means losing \$2.00. Therefore, if 2 bids \$2.80, then 1 will pass if his last bid was \$2.00 or less. In fact, any bid of \$2.10 or more will beat any last bid of \$2.00 or less for that very reason. So, whoever bids \$2.10 first establishes a credible threat to go all the way up to \$3.00 in order to win. This player has already lost \$2.10, but it is worth spending \$.90 in order to win the dollar.

Therefore, a bid of \$2.10 is "the same" as a bid of \$3.00. This means that in order to beat a \$2.00 bid, it is sufficient to bid \$2.10, nothing more. By a similar logic, \$2.00 beats all bids of \$1.10 or less. Since beating \$2.00 involves bidding \$2.10, if the player's last bid was for \$1.10 or less, there is no reason to outbid a \$2.00 bid because doing so yields at most a loss of \$1.10, which is at most equal to the player's current bid. In fact, even \$1.20 is sufficient to beat a bid of \$1.10 or less. This is because once someone bids \$1.20, it is worthwhile for that player to continue up to \$2.10 to guarantee a victory (in which case they will stand to lose \$1.10 instead of \$1.20).

Therefore, a bid of \$1.20 is "the same" as a bid of \$2.10, which is "the same" as a bid of \$3.00. By this logic, a bid of \$1.10 beats bids of 20 cents or less. Because beating \$1.10 involves bidding \$1.20, if the player's last bid was for 20 cents or less, there is no reason to win. Only a player who bids 30 cents has a credible threat to go up to \$1.20. So, any amount down to 30 cents beats a bid of 20 cents or less. If someone bids 30 cents, no one with a bid of 20 cents or less has an incentive to challenge.

Therefore, the first player to bid should bid 30 cents and the auction should end. How does that correspond to our outcome? (Probably not too well.)

Let's look at an extensive-form representation of a simpler variant, where two players have \$3.00 each but they can only bid in dollar increments in an auction for \$2.00.



We begin with the longest terminal history, (\$1, \$2), and consider 1's decision there. Since 1 is indifferent between Passing and bidding \$3, both are optimal (we do not use the indifference assumption as in the informal example). If he passes, then player 2 is indifferent between passing and bidding \$2 at the decision node following history (\$1). If, on the other hand player 1 bids \$3, then player 2's best course is to pass. So, the subgame beginning at the information set (\$1) has three SPE in pure strategies: (Pass, Pass), (\$2, Pass), and (Pass, \$3).

At the decision node following history (\$2), player 2's unique optimal action is to pass, and so the subgame perfect equilibrium there is (Pass). Therefore, player 2's strategy must specify Pass for this decision node in any SPE.

Consider now player 1's initial decision. If the players' strategies are such that they play the (Pass, Pass) SPE in the subgame after history (\$1), then player 1 does best by bidding \$1 at the outset. Therefore, ⟨(\$1, Pass), (Pass, Pass)⟩ is an SPE of the game.

If the strategies specify the SPE (\$2, Pass), then player 1's best actions are to bid \$2 or Pass at the initial node. There are two SPE: ⟨(\$2, Pass), (\$2, Pass)⟩ and ⟨(Pass, Pass), (\$2, Pass)⟩.

If the strategies specify the SPE (Pass, \$3), then player 1's best action is again to bid \$1, and so there is one other SPE: ⟨(\$1, \$3), (Pass, Pass)⟩.

The game has four subgame perfect equilibria in pure strategies. It only has three outcomes: (a) player 1 bids \$1 and player 2 passes, yielding player 1 a gain of \$1 and player 2 a payoff of \$0; (b) player 1 passes and both get \$0; (c) player 1 bids \$2 and player 2 passes, and both get \$0.

Now suppose we introduce our assumption, which states that if a player is indifferent between passing now and bidding something that will yield him the same payoff later given the other player's strategy, then the player passes. This means that in the longest history player 1 will Pass instead of bidding \$3, which further implies player 2 will Pass at the preceding node, and so this subgame will have only one SPE: ⟨Pass, Pass⟩. But now player 1 has a unique best initial action, which is to bid \$1. Therefore, the unique SPE under this assumption is ⟨(\$1, Pass), (Pass, Pass)⟩. The outcome is that player 1 bids \$1 and player 2 passes. This corresponds closely to the outcome in our discussion above.

There is a general formula that you can use to calculate the optimal size of the first bid, which depends on the amount of money available to each bidder, the size of the award, and the amount of bid increments. Let the bidding increment be represented by one unit (so the unit in our example is a dime). If each player has b units available for bidding, and the award is v units, the the optimal bid is $((b-1) \mod (v-1)) + 1$. In our example, this translates to $(30-1) \mod (10-1) + 1 = 3$ units, which equals 30 cents, just as we found.

It is interesting to note that the size of the optimal bid is very sensitive to the amount each player has available for bidding. If each player has \$2.80 instead of \$3.00, then the optimal bid is $(28-1) \mod (10-1) + 1 = 1$, or just 10 cents. If, however, each has \$2.70, then the optimal bid is $(27-1) \mod (10-1) + 1 = 9$, or 90 cents.

The Dollar Auction was first described by Martin Shubik who reported regular gains from playing the game in large undergraduate classes. The game is a very useful thought experiment about escalation. At the outset, both players are trying to win the prize by costly escalation, but at some point the escalation acquires momentum of its own and players continue paying costs to avoid paying the larger costs of capitulating. The requirement that both highest bidders pay the cost captures the idea of escalation. The general solution (the formula for the size of the optimal bids) was published by Barry O'Neill.

4.3 Multi-Stage Games with Observed Actions

There is a very special but useful class of extensive form games, which have "stages" such that (a) in each stage every player knows all the actions taken by any player at any previous stage, and (b) players move simultaneously within each stage. So, "multi-stage games with observed actions" is just a fancy name for the class of extensive form games, where the players move simultaneously within each stage and know the actions that were chosen in all past stages. Most, but not all, of the games we have seen so far fall into this category.

4.3.1 Sophisticated Voting and Agenda Control

Suppose there are three players, $\mathcal{I} = \{1, 2, 3\}$, who must choose one from three alternatives $\mathcal{X} = \{x, y, z\}$. Their preferences are as follows:

- Player 1: $x \succ y \succ z$;
- Player 2: $y \succ z \succ x$;
- Player 3: $z \succ x \succ y$.

They must make their choice through majority rule voting in a two-stage process. They first vote on two of the alternatives and the winner is then pitted against the remaining alternative in the second round. Players cast their votes simultaneously in each of the two rounds. Suppose player 2 controls the agenda—that is, she can decide which two alternatives are to be voted on in the first round. What is her choice?

We need to find the SPE. Clearly, player 2 will set an agenda that ensures that y is the winner if that's possible. We have to consider the three possible situations, depending on which two alternatives are selected in the first round. To find the subgame-perfect equilibrium, we have to ensure that the strategies are optimal in all subgames. There are three generic subgames that begin with the second round: depending on the winning alternative in the first round, the second round can involve a vote on (x, y), or (x, z), or (y, z). So let's analyze each of these subgames.

Note first that because there are only two alternatives and three players, it follows that in any PSNE, at least two players must vote for the same alternative. Observe now that casting a *sincere* vote—that is, voting for the preferred alternative—is weakly dominant for each player. If the other two players vote for different alternatives, then the third player's vote is decisive, and it is strictly better to cast it sincerely. If, on the other hand, the two other players vote for the same alternative, then the third player's vote cannot change the outcome. Therefore, there are two possible PSNE in these subgames: either all players vote sincerely or they all vote for the same alternative. Technically, this means that when x is pitted against y, it is possible to get y to win: the strategy profile in which all players vote for y is a Nash equilibrium. However, this PSNE requires two of the players to vote against their preferred alternative and *expect that they do so*, which seems highly implausible. In this instance, I would rule out the PSNE involving weakly dominated strategies. Using only weakly dominant strategies then yields a unique PSNE with sincere voting for each of the subgames, as follows:

- if (x, y), then x wins (1 and 3 vote x and 2 votes y);
- if (x, z), then z wins (2 and 3 vote z and 1 votes x);

- if (y, z), then y wins (1 and 2 vote y and 3 votes z).[9]

In other words, we know that in SPE the second round will involve sincere voting and produce a winner accordingly.

Going back to our original question, how is player 2 to set the agenda? Clearly, if y is ever to emerge as the winner given that the 2nd round will involve sincere voting, it will have to be pitted there against z because against x it will lose. Since z defeats x in sincere voting, then perhaps choosing (x, z) as the first round agenda would work?

The answer is that it will not. Suppose player 2 set the agenda with (x, z) in the first round and everyone voted sincerely. Then the winner would be z, and in the second round the winner would be y. But players can anticipate this outcome. In particular, player 3 knows that the winner of the first comparison would go on to compete with y and if y prevails in the 2nd round, he will get his worst possible alternative. Since y will beat z with sincere voting, this means that he really does not want z to win in the first round. Of votes are cast sincerely in the first round, then player 1 is voting for x while players 2 and 3 are voting for z. However, if player 3 deviated and cast a *sophisticated* vote for x instead, then x will win the first round, and in the second round the sincere vote on (x, y) would leave x as the winner. Although the sophisticated vote does not enable player 3 to get his most preferred alternative, it does enable him to avoid the worst possible one.[10]

This now means that whatever player 2 chooses, her agenda has to be invulnerable to sophisticated voting. Well, as they say, if you can't beat them, join them: player 2 will exploit the sophistication of the players by setting the agenda for the first round to (x, y). Observe that with sincere voting, x would defeat y. However, this would pit x against z in the 2nd round, in which case z will prevail. Player 1 can foresee this and since z is his worst possible alternative, he will cast a sophisticated vote for y against her preference for x over y. Doing so would ensure that y will go on to the 2nd round and defeat z, which gives him the second-best outcome. Of course, our devious player 2 can now enjoy her most preferred alternative. Therefore, the profile

$$\langle (y, xy), (y, z, y), (x, zz) \rangle$$

is a subgame-perfect equilibrium when (x, y) is the pair in the first round. The strategies are specified as a triple over the (x, y) choice in the first round, and then the (x, z) and (y, z) possible subgames in the second round. In this SPE player 1 is casting a sophisticated vote (note that player 2's manipulation pays off even though she votes sincerely). Alternative y defeats x, and then goes on to defeat z in a sincere vote in the 2nd round. Since this SPE yields player 2 her most preferred alternative, the overall SPE of the game involves her setting the agenda such that (x, y) are the two competing alternatives in the first round.

We know from McKelvey's Chaos Theorem that if players vote sincerely using majority rule to select winners in pairwise comparisons, then any outcome is possible provided no equilibrium position exists. (That is, for any two alternatives, one can always find an agenda that guarantees that one beats the other.) With sophisticated voting, this chaos is a bit reduced: for any two alternatives, there will be an agenda that guarantees that one defeats

[9] Of course, there are also the PSNE in which two players vote sincerely and the third, whose vote is irrelevant, also votes for the same alternative. For instance, in (x, y), there is PSNE in which all players vote for x. Since player 3's choice cannot affect the outcome, he might as well vote insincerely.

[10] Here, as before, there are equilibria in which all three players vote for the same alternative and two of them vote against their preferences. For instance, in (x, z) this would require them all to vote for x. Coupling this with any PSNE in the 2nd round will yield an SPE, but the solution is implausible for the same reasons we discussed already.

the other only if the winner can also beat the loser in a majority vote with sincere voting or there is a third alternative that can defeat the loser and itself be defeated by the winner in a majority vote with sincere voting (this is due to Shepsle and Weingast). In our situation y can beat z on its own with sincere voting, and y can beat x through z because z can defeat x in sincere voting and y in turn defeats z. Hence, there is an agenda that ensures y is reachable.[11]

Agenda-setting give player 2 the ability to impose her most preferred outcome and there is nothing (in this instance) that the others can do. For instance, player 1 and player 3 cannot collude to defeat her obvious intent. To see this, suppose player 3 proposed a deal to player 1: if player 1 would vote sincerely for x in the first round, then player 3 would reward him by voting for x on the 2nd round. Since x will then beat y in the first round, player 3's insincere vote in the second round would ensure that x will defeat z as well. This would benefit both players: player 1 would get his most preferred outcome and player 3 would avoid the worst outcome y and get his second-best. Unfortunately (for player 3), he cannot make a credible promise to cast an insincere vote. If x defeats y in the first round, then player 3 can get his most preferred outcome by voting sincerely in (x, z) in the second round. Therefore, he would renege on his pledge, so player 1 has no incentive to believe him. But since this reneging would saddle player 1 with his worst outcome, player 1 would strictly prefer to cast his sophisticated vote in the first round even though he is perfectly aware of how player 2 has manipulated the agenda to her advantage. The inability to make credible promises, like the inability to make credible threats, can seriously hurt players. In this instance, player 3 gets the worst of it.

4.3.2 The Ultimatum Game

Two players want to split a pie of size $\pi > 0$. Player 1 offers a division $x \in [0, \pi]$ according to which his share is x and player 2's share is $\pi - x$. If player 2 accepts this offer, the pie is divided accordingly. If player 2 rejects this offer, neither player receives anything. The extensive form of this game is represented in Fig. 18 (p. 31).

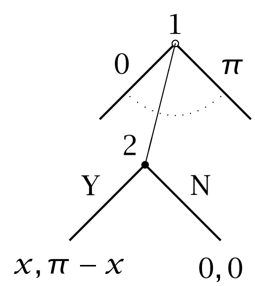


Figure 18: The Ultimatum Game.

In this game, player 1 has a continuum of action available at the initial node, while player 2 has only two actions. (The continuum of actions ranging from offering 0 to offering the entire pie is represented by the dotted curve connecting the two extremes.) When player 1 makes some offer, player 2 can only accept or reject it. There is an infinite number of subgames following a history of length 1 (i.e. following a proposal by player 1). Each history is uniquely identified by the proposal, x. In all subgames with $x < \pi$, player 2's optimal action is to

[11] In our game, each of the three outcome is possible with an appropriate agenda. Sophisticated voting does not reduce the chaos.

accept because doing so yields a strictly positive payoff which is higher than 0, which is what she would get by rejecting. In the subgame following the history $x = \pi$, however, player 2 is indifferent between accepting and rejecting. So in a subgame perfect equilibrium player 2's strategy either accepts all offers (including $x = \pi$) or accepts all offers $x < \pi$ and rejects $x = \pi$.

Given these strategies, consider player 1's optimal strategy. We have to find player 1's optimal offer for every SPE strategy of player 2. If player 2 accepts all offers, then player 1's optimal offer is $x = \pi$ because this yields the highest payoff. If player 2 rejects $x = \pi$ but accepts all other offers, *there is no optimal offer for player 1!* To see this, suppose player 1 offered some $x < \pi$, which player 2 accepts. But because player 2 accepts all $x < \pi$, player 1 can improve his payoff by offering some x' such that $x < x' < \pi$, which player 2 will also accept but which yields player 1 a strictly better payoff.

Therefore, the ultimatum game has a unique subgame perfect equilibrium, in which player 1 offers $x = \pi$ and player 2 accepts all offers. The outcome is that player 1 gets to keep the entire pie, while player 2's payoff is zero.

This one-sided result comes for two reasons. First, player 2 is not allowed to make any counteroffers. If we relaxed this assumption, the SPE will be different. (In fact, in the next section we shall analyze a very general bargaining model.) Second, the reason player 1 does not have an optimal proposal when player 2 accepts all offers has to do with him being able to always do a little better by offering to keep slightly more. Because the pie is perfectly divisible, there is nothing to pin the offers. However, making the pie discrete (e.g. by slicing it into n equal pieces and then bargaining over the number of pieces each player gets to keep) will change this as well. (You will do this in your homework.)

4.3.3 The Holdup Game

We now analyze a three-stage game. Before playing the Ultimatum Game from the previous section, player 2 can determine the size of the pie by exerting a small effort, $e_S > 0$ resulting in a small pie of size π_S, or a large effort, $e_L > e_S$, resulting in a larger pie of size $\pi_L > \pi_S$. Since player 2 hates exerting any efforts, her payoff from obtaining a share of size x is $x - e$, where e is the amount of effort expended. The extensive form of this game is presented in Fig. 19 (p. 32).

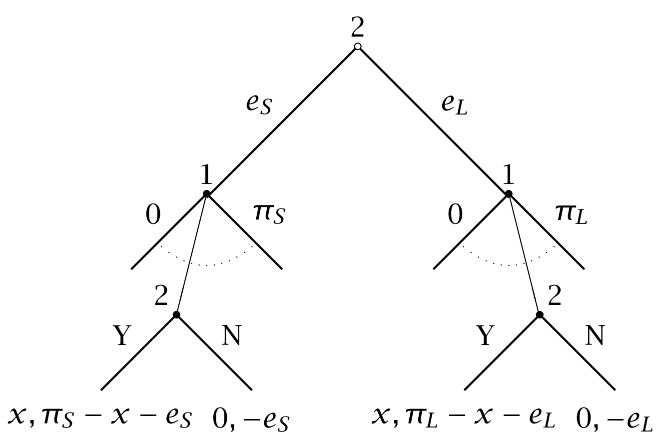


Figure 19: The Holdup Game.

We have already analyzed the Ultimatum Game, so each subgame that follows player 2's effort has a unique SPE where player 1 proposes $x = \pi$ and player 2 accepts all offers (note that the difference between this version and the one we saw above is that player 2 gets a strictly negative payoff if she rejects an offer instead of 0). So, in the subgame following e_S, player 1 offers π_S and in the subgame following e_L he offers π_L. In both cases player 2 accepts these proposals, resulting in payoffs of $-e_S$ and $-e_L$ respectively. Given these SPE strategies, player 2's optimal action at the initial node is to expend little effort, or e_S because doing so yields a strictly better payoff.

We conclude that the SPE of the Holdup Game is as follows. Player 1's strategy is (π_S, π_L) and player 2's strategy is (e_S, Y, Y), where Y means "accept all offers." The outcome of the game is that player 2 invests little effort, e_S, and player 1 obtains the entire small pie π_S.

Note that this equilibrium does not depend on the values of s_S, e_L, π_S, π_L as long as $e_S < e_L$. Even if π_L is much larger than π_S and e_L is only slightly higher than e_S, player 2 would still exert little effort in SPE although it would be better for both players if player 2 exerted e_L (remember, only *slightly* larger than e_S) and obtained a slice of the larger pie. The problem is that player 1 cannot credibly promise to give that slice tho player 2. Once player 2 expends the effort, she can be "held up" for the entire pie by player 1.

This result holds for similar games where the bargaining procedure yields a more equitable distribution. If player 2 must expend more effort to generate a larger pie and if the procedure is such that some of this surplus pie goes to the other player, then for some values of player 2's cost of exerting this effort, she would strictly prefer to exert little effort. Although there are many outcomes where both players would be strictly better off if player 2 exerted more effort, these cannot be sustained in equilibrium because of player 1's incentives. In the example above, player 1 would have liked to be able to commit credibly to offering some of the extra pie to induce player 2 to exert the larger effort. Just like the problem with non-credible threats, the problem of non-credible promises means that this cannot happen in subgame perfect equilibrium.

4.3.4 A Two-Stage Game with Several Static Equilibria

Consider the multi-stage game corresponding to two repetitions of the symmetric normal form game depicted in Fig. 20 (p. 33). In the first stage of the game, the two players simultaneously choose among their actions, observe the outcome, and then in the second stage play the static game again. The payoffs are simply the discounted average from the payoffs in each stage. That is, let p_i^1 represent player i's payoff at stage 1 and p_i^2 represent his payoff at stage 2. Then player i's payoff from the multi-stage game is $u_i = p_i^1 + \delta p_i^2$, where $\delta \in (0, 1)$ is the discount factor.

Player 1 \ Player 2	A	B	C
A	$0,0$	$3,4$	$6,0$
B	$4,3$	$0,0$	$0,0$
C	$0,6$	$0,0$	$5,5$

Figure 20: The Static Period Game.

If the game in Fig. 20 (p. 33) is played once, there are three Nash equilibria, two asymmetric ones in pure strategies: $\langle B, A\rangle$, $\langle A, B\rangle$, and one symmetric in mixed strategies with $\sigma(A) = 3/7$, and $\sigma(B) = 4/7$.

How do we find the MSNE? Suppose $\sigma_1(C) > 0$ in a MSNE. We have several cases to consider. First, suppose that $\sigma_1(B) > 0$ as well; that is, player 1 puts positive weight on both B and C (and possibly A). Since he is willing to mix, it follows that $4\sigma_2(A) = 5\sigma_2(C) \Rightarrow \sigma_2(C) = {}^4\!/_5\sigma_2(A)$. There are two ways to satisfy this requirement. Suppose $\sigma_2(C) = \sigma_2(A) = 0$, but in this case we obtain $\langle A, B\rangle$. Suppose $\sigma_2(C) > 0$, which implies $\sigma_2(A) > 0$ too; that is, player 2 must be willing to mix between A and C (and possibly B). This now implies that $3\sigma_1(B) + 6\sigma_1(C) = 5\sigma_1(C) \Rightarrow 3\sigma_1(B) + \sigma_1(C) = 0$, a contradiction because $\sigma_1(C) > 0$. Therefore, if $\sigma_1(C) > 0$, then $\sigma_1(B) = 0$ as well. Second, suppose that $\sigma_1(A) > 0$ as well; that is, player 1 puts positive weight on both A and C. Since he is willing to mix, it follows that $3\sigma_2(B) + 6\sigma_2(C) = 5\sigma_2(C) \Rightarrow 3\sigma_2(B) + \sigma_2(C) = 0$, which implies $\sigma_2(B) = \sigma_2(C) = 0$, which means $\sigma_2(A) = 1$, but this means that player 1 will not mix and we have $\langle B, A\rangle$. From all this, we conclude that $\sigma_1(C) = 0$ in MSNE. Analogously, symmetry gives us $\sigma_2(C) = 0$ too. This now reduces the game to the 2×2 variant in Fig. 21 (p. 34).

		Player 2	
		A	B
Player 1	A	0, 0	3, 4
	B	4, 3	0, 0

Figure 21: The Static Period Game after Some Equilibrium Logic..

This is now very easy to deal with. Since player 1 is willing to mix, it follows that $3\sigma_2(B) = 4\sigma_2(A)$ and since $\sigma_1(B) = 1 - \sigma_2(A)$, this gives us $\sigma_2(A) = {}^3\!/_7$. Analogously, we obtain $\sigma_1(A) = {}^3\!/_7$ and $\sigma_1(B) = {}^4\!/_7$. The last thing we need to do is check that the players will not want to use C given the mixtures (we already know this from the argument above, but it does not hurt to recall the MSNE requirement). It suffices to check for player 1: if he plays C, his payoff will be 0 given player 2's strategy of playing only A and B with positive probability, which is strictly worse than the expected payoff from either A or B. Hence, we do have our MSNE indeed. The payoffs in the three equilibria are $(4, 3)$, $(3, 4)$, and $({}^{12}\!/_7, {}^{12}\!/_7)$ respectively.

The efficient payoff $(5, 5)$ is not attainable in equilibrium if the game is played once.[12] However, consider the following strategy profile for the two-stage game:

- Player 1: play C at the first stage. If the outcome is (C, C), play B at the second stage, otherwise play $\sigma_1(A) = {}^3\!/_7, \sigma_1(B) = {}^4\!/_7$ at the second stage;

- Player 2: play C at the first stage. If the outcome is (C, C), play A at the second stage, otherwise play $\sigma_1(A) = {}^3\!/_7, \sigma_2(B) = {}^4\!/_7$ at the second stage.

Is it subgame perfect? Since the strategies at the second stage specify playing Nash equilibrium profiles for all possible second stages, the strategies are optimal there. At the first stage players can deviate and increase their payoffs by 1 from 5 to 6 (either player can choose A). However, doing so results in playing the mixed strategy Nash equilibrium at the second stage, which lowers their payoffs to ${}^{12}\!/_7$ from 4 for player 1 and from 3 for player 2. Thus,

[12] An outcome is *efficient* if it is not possible to make some player better off without making the other one worse off. The outcomes with payoffs $(0, 0)$ are all inefficient, as are the outcomes with payoffs $(4, 3)$ and $(3, 4)$. However, the outcomes $(6, 0)$ and $(0, 6)$ are also efficient.

player 1 will not deviate if:

$$\begin{aligned} 6 + \delta\,(12/7) &\le 5 + \delta(4) \\ 1 &\le \delta\,(4 - 12/7) \\ \delta &\ge 7/16 \end{aligned}$$

Similarly, player 2 will not deviate if:

$$\begin{aligned} 6 + \delta\,(12/7) &\le 5 + \delta(3) \\ 1 &\le \delta\,(3 - 12/7) \\ \delta &\ge 7/9 \end{aligned}$$

We conclude that the strategy profile specified above is a subgame perfect equilibrium if $\delta \ge 7/9$. In effect, players can attain the non-Nash efficient outcome at stage 1 by threatening to revert to the worst possible Nash equilibrium at stage 2. This technique will be very useful when analyzing infinitely repeated games, where we shall see analogous results.

4.4 The Principle of Optimality

This is an important result that you will make heavy use of both for multi-stage games and for infinitely repeated games that we shall look at next time. The principle states that to check whether a strategy profile of a multi-stage game with observed actions is subgame perfect, it suffices to check whether there are any histories h^t where some player i can profit by deviating only from the actions prescribed by $s_i(h^t)$ and conforming to s_i thereafter. In other words, for games with arbitrarily long (but finite) histories,[13] it suffices to check if some player can profit by deviating only at a single point in the game and then continuing to play his equilibrium strategy. That is, we do not have to check deviations that involve actions at several points in the game. You should be able to see how this simplifies matters considerably.

The following theorem is variously called "The One-Shot (or One-Stage) Deviation Principle," and is essentially the principle of optimality in dynamic programming. Since this is such a nice result and because it may not be obvious why it holds, we shall go through the proof.

Theorem 6 (One-Shot Deviation Principle for Finite Horizon Games). *In a finite multi-stage game with observed actions, strategy profile (s_i^*, s_{-i}^*) is a subgame perfect equilibrium if, and only if, it satisfies the condition that no player i can gain by deviating from s_i^* in a single stage and conforming to s_i^* thereafter while all other players stick to s_{-i}^*.*

Proof. (Necessity.) This follows immediately from the definition of SPE. If (s_i^*, s_{-i}^*) is subgame perfect equilibrium, then no player has an incentive to deviate in any subgame.[14]

(Sufficiency.) Suppose that (s_i^*, s_{-i}^*) satisfies the one-shot deviation principle but is not subgame perfect. This means that there is a subgame after some history h such that there is another strategy, $s_i \neq s_i^*$ that is a better response to s_{-i}^* than s_i^* is in the subgame starting

[13] We shall see that this principle also works for infinitely repeated games under some conditions that will always be met by the games we consider.

[14] This does *not* hold for Nash equilibrium, which may prescribe suboptimal actions off the equilibrium path (i.e. in some subgames).

with h. Let $\hat{t}$ be the largest t such that, for some h^t, $s_i(h^t) \neq s_i^*(h^t)$. (That is, $h^{\hat{t}}$ is the history that includes all deviations.) Because s_i^* satisfies the OSDP, $h^{\hat{t}}$ is longer than h and, because the game is finite, $h^{\hat{t}}$ is finite as well. Now consider an alternate strategy $\hat{s}_i$ that agrees with s_i at all $t < \hat{t}$ and follows s_i^* from stage $\hat{t}$ on. Because $\hat{s}_i$ is the same as s_i^* in the subgame beginning with $h^{\hat{t}+1}$ and the same as s_i in all subgames with $t < \hat{t}$, the OSDP implies that it is as good a response to s_{-i}^* as s_i in every subgame starting at t with history h^t. If $\hat{t} = t + 1$, then $\hat{s}_1 = s_1^*$, which contradicts the hypothesis that s_1 improves on s_1^*. If $\hat{t} > t + 1$, construct a strategy that agrees with s_1 until $t - 2$, and argue that it is as good a response as s_1, and so on. The sequence of improving deviations unravels from its endpoint. □

The proof works as follows. You start from the last deviation in a sequence of multiple deviations and argue that it cannot be profitable by itself, or else the OSDP would be violated. This now means that if you use the multiple-deviation strategy up to that point and follow the original OSDP strategy from that point on, you would get at least as good a payoff (again, because the last deviation could not have been the profitable one, so the original OSDP strategy will do at least as good in that subgame). You then go up one step to the new "last" deviation and argue that this deviation cannot be profitable either: since we are comparing a subgame with this deviation and the original OSDP strategy to follow with the OSDP strategy itself, the fact that the original strategy satisfies OSDP implies that this particular deviation cannot be profitable. Hence, we can replace this deviation with the action from the OSDP strategy too and obtain at least as good a payoff as the multi-deviation strategy. You repeat this process until you reach the first stage with a deviation and you reach the contradiction because this deviation cannot be profitable by itself either. In other words, if a strategy satisfies OSDP, it must be subgame perfect.

An example here may be helpful. Since in equilibrium we hold all other players strategies constant when we check for profitable deviations, the diagram in Fig. 22 (p. 36) omits the strategies for the other players and shows only player 1's moves at his information sets. Label the information sets consecutively with small Roman numerals for ease of exposition. Suppose that the strategy $(adegi)$ satisfies OSDP. We want to show that there will be no more profitable other strategies even if they involve multiple deviations from this one. To make the illustration even more helpful, I have bolded the actions specified by the OSDP strategy.

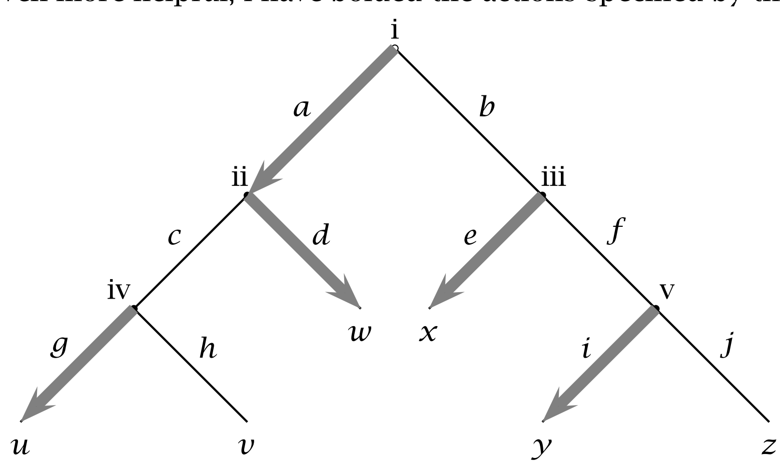


Figure 22: An Illustration with $s = (adegi)$ Satisfying OSDP.

Because $(adegi)$ satisfies OSDP, we can infer certain things about the ordering of the payoffs. For example, OSDP implies that changing from g to h at (iv) cannot be profitable, which implies $u \geq v$. Also, at (v), changing from i to j cannot be profitable, so $y \geq z$. At (ii), changing to c cannot be profitable; since the strategy specifies playing g at (iv), this deviation leads to $w \geq u$. At (iii), changing to f cannot be profitable. Since the original strategy specifies i at (v), this deviation will lead to $x \geq y$. Finally, at (i) changing to b cannot be profitable. Since the original strategy specifies e at (iii), this deviation will lead to $w \geq x$. The implications of OSDP are listed as follows:

$$\text{at (i)} : w \geq x \tag{1}$$
$$\text{at (ii)} : w \geq u \tag{2}$$
$$\text{at (iii)} : x \geq y \tag{3}$$
$$\text{at (iv)} : u \geq v \tag{4}$$
$$\text{at (v)} : y \geq z \tag{5}$$

These inequalities now imply some further relationships: from the first and the third, we get $w \geq y$, and putting this together with the last yields $w \geq z$ as well. Furthermore, from the third and last we obtain $x \geq z$, and from the second and fourth we obtain $w \geq v$. Putting everything together yields the following orderings of the payoffs:

$$w \geq x \geq y \geq z \qquad \text{and} \qquad w \geq u \geq v.$$

We can now check whether there exist any profitable multi-stage deviations. (Obviously, there will be no single-stage profitable deviations because the strategy satisfies OSDP.) Take, for example, an alternative strategy that deviates at (ii) and (iv); that is in the subgame starting at (ii), it specifies ch. This will lead to the outcome v, which cannot improve on w, the outcome from following the original strategy. Consider another alternative strategy which deviates twice in the subgame starting at (iii); i.e., it prescribes fj, which would lead to the outcome z. This cannot improve on x, the outcome the player would get from following the original strategy. Going up to (i), consider a strategy that deviates at (i) and (v). That is, it prescribes b and j at these information sets. Since (v) is still never reached, this actually boils down to a one-shot deviation with the outcome x, which (not surprisingly) cannot improve on w, which is what the player can get from following the original strategy. What if he deviated at (i) and (iii) instead? This would lead to y, which is also no better than w. What if he deviated at (i), (iii), and (v)? This would lead to z, which is also no better than w. Since all other deviations that start at (i) leave (ii) and (iv) off the path of play, there is no need to consider them. This example then shows how OSDP implies subgame-perfection. Intuitively, if a strategy satisfies OSDP, then it implies a certain preference ordering, which in turn ensures that no multi-stage deviations will be profitable.

To see how the proof would work here. Take the longest deviation, e.g., a strategy that deviates at (i), (iii), and (v). Since it leaves (ii) and (iv) off the path, let's consider $(bdfgj)$ as such a supposedly better alternative. Observe now that because $(adegi)$ satisfies OSDP, the deviation to j at (v) cannot be improving. This means that the strategy $(bdfgi)$ is at least as good as $(bdfgj)$. Hence, if $(bdfgj)$ is better than the original, then $(bdfgi)$ must also be better. Consider now $(bdfgi)$: since it matches the original at (v), OSDP implies that the deviation to f cannot be improving. Hence, the strategy $(bdegi)$ is at least as good as $(bdfgi)$, which implies it is also at least as good as $(bdfgj)$. Hence, if $(bdfgj)$ is better than the original, then $(bdegi)$ must also be better. However, $(bdegi)$ matches the original strategy at all information sets except (i); i.e., it involves a one-shot deviation to b which cannot be improving by OSDP. Since $(bdegi)$ cannot improve on $(adegi)$, neither can $(bdfgj)$, a contradiction to the supposition that it is better than $(adegi)$. This is essentially how the proof works.

Let's now see what OSDP gets you. Consider the game in Fig. 23 (p. 38). The SPE, which you can obtain by backward induction, is $((bf), d)$, with the outcome $(3, 3)$.

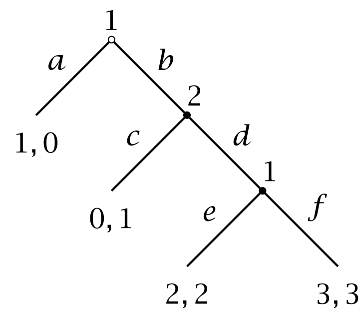


Figure 23: The One-Shot Deviation Principle.

Let's now check if player 1's strategy is indeed subgame perfect. Recall that this requires that it is optimal for all subgames. This is easy to see for the subgame that begins with player 1's second information set that follows history (b, d). How about player 1's choice at the first information set? If we were to examine all possible deviations, we must check the alternative strategies (ae), (af), and (be) because these are the other strategies available for that subgame. The one-shot deviation principle allows us to check just one thing: whether player 1 can benefit by deviating from b to a at his first information set! (This is why the OSDP is your friend.) In this case, deviating to a would get player 1 a payoff of 1 instead of 3, which he would get if he stuck to his equilibrium strategy. Therefore, this deviation is not profitable. We already saw that deviating to e in the subgame that begins with player 1's second information set is not profitable either. Therefore, by the OSDP, the strategy is subgame perfect.

This proof does not work for games with infinite horizon because the essential step in it requires that there is a finite number of possible deviations. However, in a game with infinite horizon, there are strategies with an infinite number of deviations. Fortunately, if we assume that the payoff function takes the form of a discounted sum of per-period payoffs, we can get the result "back." Again fortunately, all infinite-horizon games that we shall look at will have payoff functions that meet this requirement. The OSDP may not be that helpful in finding SPE. However, it's extremely useful for checking whether some profile is SPE.

4.5 The Rubinstein Bargaining Model

There are at least two basic ways one can approach the bargaining problem. (The bargaining problem refers to how people would divide some finite benefit among themselves.) Nash initiated the axiomatic approach with his Nash Bargaining Solution (he did not call it that, of course). This involves postulating some desirable characteristics that the distribution must meet and then determining whether there is a solution that meets these requirements. This is very prominent in economics but we shall not deal with it here.

Instead, we shall look at strategic bargaining. Unlike the axiomatic solution, this approach involves specifying the bargaining protocol (i.e. who gets to make offers, who gets to respond to offers, and when) and then solving the resulting extensive form game.

People began analyzing simple two-stage games (e.g. ultimatum game where one player makes an offer and the other gets to accept or reject it) to gain insight into the dynamics of bargaining. Slowly they moved to more complicated settings where one player makes all the offers while the other accepts or rejects, with no limit to the number of offers that can be made. The most attractive protocol is the alternating-offers protocol where players take turns making offers and responding to the other player's last offer.

The alternating-offers game was made famous by Ariel Rubinstein in 1982 when he published a paper showing that while this game has infinitely many Nash equilibria (with any division supportable in equilibrium), it had a unique subgame perfect equilibrium! Now this is a great result and since it is the foundation of most contemporary literature on strategic bargaining, we shall explore it in some detail.[15]

4.5.1 The Basic Alternating-Offers Model

Two players, A and B, bargain over a partition of a pie of size $\pi > 0$ according to the following procedure. At time $t = 0$ player A makes an offer to player B about how to partition the pie. If player B accepts the offer, then an agreement is made and they divide the pie accordingly, ending the game. If player B rejects the offer, then she makes a counteroffer at time $t = 1$. If the counteroffer is accepted by player A, the players divide the pie accordingly and the game ends. If player A rejects the offer, then he makes a counter-counteroffer at time $t = 2$. This process of alternating offers and counteroffers continues until some player accepts an offer.

To make the above a little more precise, we describe the model formally. Two players, A and B, make offers at discrete points in time indexed by $t = (0, 1, 2, \ldots)$. At time t when t is even (i.e. $t = 0, 2, 4, \ldots$) player A offers $x \in [0, \pi]$ where x is the share of the pie A would keep and $\pi - x$ is the share B would keep in case of an agreement. If B accepts the offer, the division of the pie is $(x, \pi - x)$. If player B rejects the offer, then at time $t + 1$ she makes a counteroffer $y \in [0, \pi]$. If player A accepts the offer, the division $(\pi - y, y)$ obtains. Generally, we shall specify a proposal as an ordered pair, with the first number representing player A's share. Since this share uniquely determines player B's share (and vice versa) each proposal can be uniquely characterized by the share the proposer offers to keep for himself.

The payoffs are as follows. While players disagree, neither receives anything (which means that if they perpetually disagree then each player's payoff is zero). If some player agrees on a partition $(x, \pi - x)$ at some time t, player A's payoff is $\delta^t x$ and player B's payoff is $\delta^t(\pi - x)$.

The players discount the future with a common discount factor $\delta \in (0, 1)$. This captures the time preferences of the players: it is better to obtain an agreement sooner than later. Here's how it works. Suppose agreement is reached on a partition $(2, \pi - 2)$ at some time t. Player A's payoff is $2\delta^t$. Since $0 < \delta < 1$, as t increases, δ^t decreases. Let $\delta = .9$ (so a dollar tomorrow is only worth 90 cents today). If this agreement is reached at $t = 0$, player A's payoff $(2)(.9)^0 = 2$. If the agreement is reached at $t = 1$, player A's payoff is $(2)(.9)^1 = 1.8$. If it happens at $t = 2$, player A's payoff is $(2)(.9)^2 = 1.62$. At $t = 10$, the payoff is $(2)(.9)^{10} \approx .697$, at $t = 100$, it is only $(2)(.9)^{100} = .000053$, and so on and so forth. The point is that the further in the future a player gets some share, the less attractive this same share is compared to getting it sooner.

[15] The Rubinstein bargaining model is extremely attractive because it can be easily modified, adapted, and extended to various settings. There is significant cottage industry that does just that. The Muthoo (1999) book gives an excellent overview of the most important developments. The discussion that follows is taken almost verbatim from Muthoo's book. If you are serious about studying bargaining, you should definitely get this book. Yours truly also tends to use variations of the Rubinstein model in his own work on intrawar negotiations.

This completes the formal description of the game. You can draw the extensive form tree for several periods, but since the game is not finite (there's an infinite number of possible offers at each information set *and* the longest terminal history is infinite—the one where players always reject offers), we cannot draw the entire tree.

4.5.2 Nash Equilibria

Let's find the Nash equilibria in pure strategies for this game. Actually, we cannot find all Nash equilibria because there's an infinite number of those. What we can do, however, is characterize the payoffs that players can get in equilibrium. It turns out that *any division of the pie* can be supported in some Nash equilibrium. To see this, consider the strategies where player A demands $x \in (0, \pi)$ in the first period, then π in each subsequent period where he gets to make an offer, and always rejects all offers. This is a valid strategy for the bargaining game. Given this strategy, player B does strictly better by accepting x instead of rejecting forever, so she accepts the initial offer and rejects all subsequent offers. Given that B accepts the offer, player A's strategy is optimal.

The problem, of course, is that Nash equilibrium requires strategies to be mutually best responses only along the equilibrium path. It is just not reasonable to suppose that player A can credibly commit to rejecting all offers regardless of what player B does. To see this, suppose at some time $t > 0$, player B offers $y < \pi$ to player A. According to the Nash equilibrium strategy, player A would reject this (and all subsequent offers) which yields a payoff of 0. But player A can do strictly better by accepting $\pi - y > 0$! The Nash equilibrium is not subgame perfect because player A cannot credibly threaten to reject all offers.

4.5.3 The Unique Subgame Perfect Equilibrium

Since this is an infinite horizon game, we cannot use backward induction to solve it. However, since every subgame that begins with an offer by some player is structurally identical with all subgames that begin with an offer by that player, we shall look for an equilibrium with two intuitive properties: (1) *no delay:* whenever a player has to make an offer, the equilibrium offer is immediately accepted by the other player; and (2) *stationarity:* in equilibrium, a player always makes the same offer.

It is important to realize that at this point I do not claim that such equilibrium exists—we shall look for one that has these properties. Also, I do not claim that if it does exist, it is the unique SPE of the game. We shall prove this later. However, given that the subgames are structurally identical, there is no *a priori* reason to think that offers must be non-stationary, and, if this is the case, that there should be any reason to delay agreement (given that doing so is costly). So it makes sense to look for an SPE with these properties.

Let x^* denote player A's equilibrium offer and y^* denote player B's equilibrium offer (again, because of stationarity, there is only one such offer). Consider now some arbitrary time t at which player A has to make an offer to player B. From the two properties, it follows that if B rejects the offer, she will then offer y^* in the next period (stationarity), which A will accept (no delay). So, B's payoff to rejecting A's offer is δy^*. Subgame perfection requires that B reject any offer $\pi - x < \delta y^*$ and accept any offer $\pi - x > \delta y^*$. From the no delay property, this implies $\pi - x^* \geq \delta y^*$. However, it cannot be the case that $\pi - x^* > \delta y^*$ because player A could increase his payoff by offering some x such that $\pi - x^* > \pi - x > \delta y^*$. Hence:

$$\pi - x^* = \delta y^* \tag{6}$$

Equation 6 states that in equilibrium, player B must be indifferent between accepting and rejecting player A's equilibrium offer. By a symmetric argument it follows that in equilibrium, player A must be indifferent between accepting and rejecting player B's equilibrium offer:

$$\pi - y^* = \delta x^* \tag{7}$$

Equations (6) and (7) have a unique solution:

$$x^* = y^* = \frac{\pi}{1+\delta}$$

which means that there may be at most one SPE satisfying the no delay and stationarity properties. The following proposition specifies this SPE.

PROPOSITION 1. *The following pair of strategies is a subgame perfect equilibrium of the alternating-offers game:*

- *player A always offers $x^* = \pi/(1+\delta)$ and always accepts offers $y \leq y^*$,*
- *player B always offers $y^* = \pi/(1+\delta)$ and always accepts offers $x \leq x^*$.*

Proof. We show that player A's strategy as specified in the proposition is optimal given player B's strategy. Consider an arbitrary period t where player A has to make an offer. If he follows the equilibrium strategy, the payoff is x^*. If he deviates and offers $x < x^*$, player B would accept, leaving A strictly worse off. Therefore, such deviation is not profitable. If he instead deviates by offering $x > x^*$, then player B would reject. Since player B always rejects such offers and never offers more than y^*, the best that player A can hope for in this case is $\max\{\delta(\pi - y^*), \delta^2 x^*\}$. That is, either he accepts player B's offer in the next period or rejects it and A's offer in the period after the next one is accepted. (Anything further down the road will be worse because of discounting.) However, $\delta^2 x^* < x^*$ and also $\delta(\pi - y^*) = \delta x^* < x^*$, so such deviation is not profitable. Therefore, by the one-shot deviation principle, player A's proposal rule is optimal given B's strategy.

Consider now player A's acceptance rule. At some arbitrary time t player A must decide how to respond to an offer made by player B. From the above argument we know that player A's optimal proposal is to offer x^*, which implies that it is optimal to accept an offer y if and only if $\pi - y \geq \delta x^*$. Solving this inequality yields $y \leq \pi - \delta x^*$ and substituting for x^* yields $y \leq y^*$, just as the proposition claims.

This establishes the optimality of player A's strategy. By a symmetric argument, we can show the optimality of player B's strategy. Given that these strategies are mutually best responses at any point in the game, they constitute a subgame perfect equilibrium. □

This is good but so far we have only proven that there is a unique SPE that satisfies the no delay and stationarity properties. We have not shown that there are no other subgame perfect equilibria in this game. The following proposition, whose proof involves knowing some (not much) real analysis, states this result.[16]

PROPOSITION 2. *The subgame perfect equilibrium described in Proposition 1 is the unique subgame perfect equilibrium of the alternating-offers game.*

[16] If you know what a supremum of a set is, you can ask me and I will tell you how the proposition can be proved. There is a very elegant proof due to Shaked and Sutton that is much easier to follow than the extremely complicated original proof by Rubinstein.

We are now in game theory heaven! The rather complicated-looking bargaining game has a unique SPE in which agreement is reached immediately. Player A offers x^* at $t = 0$ and player B immediately accepts this offer. The shares obtained by player A and player B in the unique equilibrium are $x^* = \pi/(1+\delta)$ and $\pi - x^* = \delta\pi/(1+\delta)$ respectively.

In equilibrium, the share depends on the discount factor and player A's equilibrium share is strictly greater than player B's equilibrium share. In this game there exists a "first-mover" advantage because player A is able to extract all the surplus from what B must forego if she rejects the initial proposal. In your homework you will be asked to find the unique stationary no delay SPE when the two players use different discount factors.

The Rubinstein bargaining model makes an important contribution to the study of negotiations. First, the stylized representation captures characteristics of most real-life negotiations: (a) players attempt to reach an agreement by making offers and counteroffers, and (b) bargaining imposes costs on both players.

Some people may argue that the infinite horizon assumption is implausible because players have finite lives. However, this involves a misunderstanding of what the infinite time horizon really represents. Rather than modeling a reality where bargaining can continue forever, it models a reality where players do not stop bargaining after some exogenously given predefined time limit. The finite horizon assumption would have the two players to stop bargaining even though each would prefer to continue doing so if agreement has not been reached. Unless there's a good explanation of who or what prevents them from continuing to bargain, the infinite horizon assumption is appropriate. (There are other good reasons to use the assumption and they have to do with the speed with which offers can be made. There are some interesting models that explore the bargaining model in the context of deadlines for reaching an agreement. All this is very neat stuff and you are strongly encouraged to read it.)

4.6 Bargaining with Fixed Costs

Osborne and Rubinstein also study an alternative specification of the alternating-offers bargaining game where delay costs are modeled not as time preferences but as direct per-period costs. These models do not behave nearly as nicely as the one we studied here, and they have not achieved widespread use in the literature.

As before, there are two players who bargain using the alternating-offers protocol with time periods indexed by t, $(t = 0, 1, 2, \ldots)$. Instead of discounting future payoffs, they pay per-period costs of delay, $c_2 > c_1 > 0$. That is, if agreement is reached at time t on $(x, \pi - x)$, then player 1's payoff is $x - tc_1$ and player 2's payoff is $\pi - x - tc_2$.

Let's look for a stationary no-delay SPE as before. Consider a period t in which player 1 makes a proposal. If player 2 rejects, then she can obtain $y^* - (t+1)c_2$ by our assumptions. If he accepts, on the other hand, she gets $\pi - x - tc_2$ because of the t period delay. Hence, player 2 will accept any $\pi - x - tc_2 \geq y^* - (t+1)c_2$, or $\pi - x \geq y^* - c_2$. To find now the maximum she can expect to demand, note that by rejecting her offer in $t+1$, player 1 will get $x^* - (t+2)c_1$ and by accepting it, he will get $\pi - y - (t+1)c_1$ because of the $t+1$ period delay up to his acceptance. Therefore, he will accept any $\pi - y - (t+1)c_1 \geq x^* - (t+2)c_1$, which reduces to $\pi - y \geq x^* - c_1$. Since player 2 will be demanding the most that player 1 will accept, it follows that $y^* = \pi - x^* + c_1$. This now means that player 2 cannot credibly commit to reject any t period offer that satisfies:

$$\pi - x \geq \pi - x^* + c_1 - c_2 \Leftrightarrow x^* - x \geq c_1 - c_2.$$

Observe now that since $c_1 < c_2$, it follows that the RHS of the second inequality is negative. Suppose now that $x^* < \pi$, then it is always possible to find $x > x^*$ such that $0 > x^* - x \geq c_1 - c_2$. For instance, taking $x = x^* - (c_1 - c_2) = x^* + (c_2 - c_1) > x^*$ because $c_2 > c_1$. Therefore, if $x^* < \pi$, it is possible to find $x > x^*$ such that player 1 will prefer to propose x instead of x^*, which contradicts the stationarity assumption. Therefore, $x^* = \pi$. This now pins down $y^* = c_1$. This yields the following result.

Proposition 3. *The following pair of strategies constitutes the unique stationary no-delay subgame perfect equilibrium in the alternating-offers bargaining game with per-period costs of delay $c_2 > c_1 > 0$:*

- *player 1 always offers $x^* = \pi$ and always accepts offers $y \leq c_1$;*
- *player 2 always offers $y^* = c_1$ and always accepts offers $x \leq \pi$.*

The SPE outcome is that player 1 grabs the entire pie in the first period.

Obviously, if $c_1 > c_2 > 0$ instead, then player 1 will get c_2 in the first period and the rest will go to player 2. In other words, the player with the lower cost of delay extracts the entire bargaining surplus, which in this case is heavily asymmetric. If the low-cost player gets to make the first offer, he will obtain the entire pie. It turns out that this SPE is also the unique SPE (if $c_1 = c_2$, then there can be multiple SPE, including some with delay).

This model is not well-behaved in the following sense. First, no matter how small the cost discrepancy is, the player with the lower cost gets everything. That is, it could be that player 1's cost is $c_1 = c_2 - \epsilon$, where $\epsilon > 0$ is arbitrarily small. Still, in the unique SPE, he obtains the entire pie. The solution is totally insensitive to the cardinal difference in the costs, only to their ordinal ranking. Note now that if the costs are very close to each other and we tweak them ever so slightly such that $c_1 > c_2$, then player 2 will get $\pi - c_2$; i.e., the prediction is totally reversed! This is not something you want in your models. It is perhaps for this reason that the fixed-cost bargaining model has not found wide acceptance as a workhorse model.

5 Critiques of Backward Induction and Subgame Perfection

Although backward induction and subgame perfection give compelling arguments for reasonable play in simple two-stage games of perfect information, things get uglier once we consider games with many players or games where each player moves several times.

5.1 Critiques of Backward Induction

There are two criticisms of BI and both have to do with questions about reasonable behavior. In my mind, the second critique has more bite than the first one, but I will give you both.

First, consider a game with n players that has the structure depicted in Fig. 24 (p. 44). Since this is a game of perfect information, we can apply the backward induction algorithm. The unique equilibrium is the profile where each player chooses C and in the outcome each player gets 2.

People have argued that this is unreasonable because in order to get the payoff of 2, all $n-1$ players must choose C. If the probability that any player chooses C is $p < 1$, independent of the others, then the probability that all $n - 1$ will choose C is p^{n-1}, which can be quite small if n is large even if p itself is very close to 1. For example, with $p = .999$ and $n = 1001$,

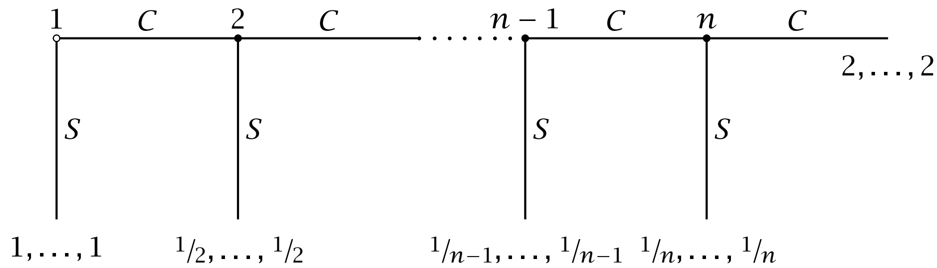


Figure 24: A Game with Many Players.

this probability is $(.999)^{1000} \approx .37$, and with $n = 10{,}001$, it is barely .00005. Moreover, player 1 has to worry that player 2 might have these concerns and might choose S in order to safeguard either against possible "mistakes" by other players in the future or the possibility that player 3, having these same concerns, might intentionally play S.

In order for the equilibrium to work, not only must players not make mistakes, but they also must know that everyone else knows the payoffs, and knows that everyone else knows the payoffs, and knows that everyone else knows that everyone else knows the payoffs, and so on and so forth. This is the *common knowledge* assumption that we've seen before. In game theory it is usually assumed that payoffs are common knowledge and so we can use arbitrarily long chains in our solutions. However, some people feel that the longer these chains, the less compelling the solution that requires them.

The second critique with BI has to do with games where the same player has to move many times. It seems to me that this is a more serious problem. Consider the game depicted in Fig. 25 (p. 44).

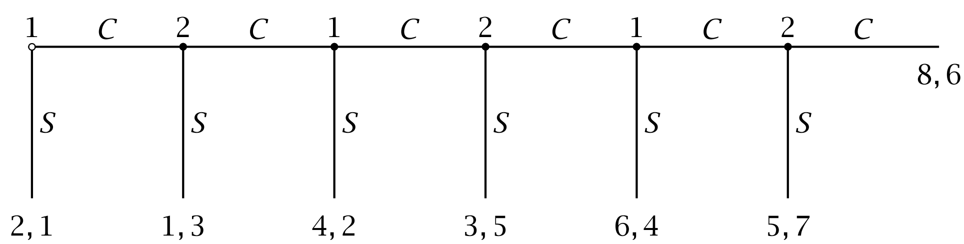


Figure 25: The Centipede Game.

The backward-induction solution is that players choose S at every information set. However, suppose that contrary to expectations player 1 chooses C at the initial node. What should player 2 do? The backward-induction solution says to play S, because player 1 will play S given a chance. However, player 1 should have played S at the initial node but did not. Since player 2's optimal behavior depends on her beliefs about player 1's behavior in the future, how does she form these beliefs following a 0-probability event? For example, if she believes that player 1 will stop with probability less than $2/3$, then she should play C because doing so will get her at least 3, which is the best she obtains from stopping.

How does player 2 form these beliefs and what beliefs are reasonable? There are two ways to address this problem. First, we may introduce some payoff uncertainty and interpret deviations from expected play by the payoffs differing from those originally thought to be most likely. Instead of conditioning beliefs on probability-0 events, this approach conditions them payoffs that are most likely given the "deviation".

Second, we may interpret the extensive form game as implicitly including the possibility that players sometimes make small "mistakes" or "trembles" whenever they act. If the probabilities of "trembles" are independent across different information sets, then no matter how often past play has failed to conform to the predictions of backward induction, a player is still justified in continuing to use backward induction for the rest of the game. There is a "trembling-hand perfect" equilibrium due to Selten that formalizes this idea. (This is a defense of backward induction.)

The question now becomes one of choosing between two possible interpretations of deviations. In Fig. 25 (p. 44), if player 2 observes C, will she interpret this as a small "mistake" by player 1 or as a signal that player 1 will choose C if given a chance? Who knows? I am more inclined toward the latter interpretation but your mileage may vary. To see why it may make sense to treat deviations as a signal, suppose we extend the centipede game to 40 periods and now suppose we find ourselves in period 20; that is, both players have played C 10 times. Is it reasonable to suppose these were all mistakes? Or that perhaps players are trying to get closer to the endgame where they would get better payoffs? In experimental settings, players usually do continue for a while although they do tend to stop well short of the end. One way we can think about this is that the game is not actually capturing everything about the players. In particular, in experiments a player may doubt the rationality of the opponent (so he may expect her to continue) or he may believe she doubts his own rationality (so she expects him to continue, which in turn makes him expect her to continue as well). At any rate, small doubts like this may move the play beyond the game-stopping first choice by player 1. This does not mean that backward induction is "wrong." What it does mean is that the full information common knowledge assumptions behind it may not be captured in experiments where real people play the Centipede Game. My reaction to this is not to abandon backward induction but to modify the model and ask: what will happen if players with small doubts about each other's rationality play the Centipede Game? This is a topic for another discussion, though.

5.2 Critiques of Subgame Perfection

Obviously, all the critiques of backward induction apply here as well. However, in addition to these problems, SP also requires that players agree on the play in a subgame even when BI cannot predict the play.

The coordination game between player 1 and 3 has three Nash equilibria: two in pure strategies with payoffs $(7, 10, 7)$, and one in mixed strategies with payoffs $(3.5, 5, 3.5)$.[17] If we specify an equilibrium in which player 1 and 3 successfully coordinate, then player 2 will choose R, and so player 1 will choose R as well, expecting a payoff of 7. If we specify the MSNE, then player 2 will choose L because R yields an expected payoff of 5 (coordination will fail half of the time). Again player 1 will choose R expecting a payoff of 8. Thus, in all SPE of this game player 1 chooses R.

Suppose, however, player 1 did not see a way to coordinate in the third stage, and hence expected a payoff of 3.5 conditional on this stage being reached, but feared that player 2 would believe that the play in the third stage would result in coordination on an efficient equilibrium. (This is not unreasonable since the two pure strategy Nash equilibria there are the efficient ones.) If player 2 had such expectations, then she would choose R, which means that player 1 would go L at the initial node!

[17] In this MSNE, each player chooses A with probability $1/2$, as you should readily see.

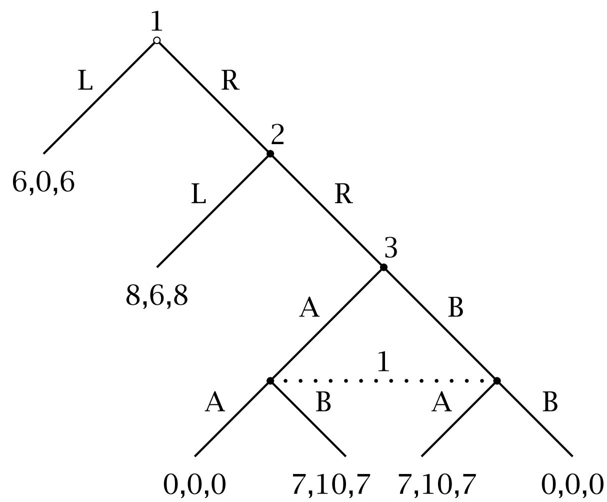


Figure 26: The Coordination Problem Game.

The problem with SPE is that *all players must expect the same Nash equilibria in all subgames.* So, while this was not a big problem for subgames with unique Nash equilibria, the critique has significant bite in cases like the one just shown. Is such a common expectation reasonable? Who knows? (It depends on the reason the equilibrium arises in the first place, which is not something we can say a whole lot about yet.)