[SQUEAKING]

[RUSTLING]

[CLICKING]

IAN BELL: All right, let's get started. So today we're going to discuss subgame perfect Nash equilibrium. And I think the best way to motivate this is to start with an example. So let's start with a very simple example of a game. Let's draw it like this.

The story is that we have two players. And the first player can choose whether to exit or enter, let's say, some market. We can think of them as firms. If they exit, then both players just get a payoff of 0. But if they enter, then the other firm, too, can choose whether to accommodate this new entrant or to fight.

So if player 1 enters, player 2 chooses accommodate or fight. If they accommodate the new entrant, that's good for both players. They both get one. If they fight against the new entrant, that's bad for both players. They both get negative 1.

So first, let's just represent this in strategic form. This is a bit of a review here. So if we represent this in strategic form, player 1 has two strategies, N and X. And player 2 has two strategies, A and F, and we can put them here.

And then we just need to fill in the payoffs. Well, if player 1 does not enter, if they exit, then both players get a payoff of 0, so we have 0, 0 0, 0. If player 1 enters, then it depends what player 2 does. And if player 1 accommodates, we get 1,1. And if player 1 fights, we get negative 1, negative 1.

So now as usual, let's try to solve for the Nash equilibria of this game. And I'll say, given, if player 2 accommodates, then I'm better off entering. If player 2 fights, then I'm better off exiting. And then if we do it the other way, we're going to get this.

So here we see that we get two Nash equilibria. Namely, our two Nash equilibria are NA and XF. So what's the story here? One Nash equilibrium is player 1 enters, anticipating that player 2 is going to accommodate their entry. The other Nash equilibrium is XF. Player 1 exits. They don't enter because they anticipate that player 2 will fight. And both of these are Nash equilibria. Both players are playing optimally given the way the other player is behaving.

But we argued last week, that in games like this, one of these is more compelling than the other. Do you find a problem with any-- does any of these Nash equilibria involve a non-credible threat that maybe makes it less compelling than the other Nash equilibrium? Yeah?

AUDIENCE: Is the player 2 will fight [INAUDIBLE]?

IAN BELL: Right, so this is a Nash equilibrium, but this involves a non-credible threat. Why? Well, in this equilibrium, player 1 is exiting. They're choosing not to enter because they anticipate that player 2 will fight them. But that's not a credible threat. If they actually were to enter, then player 2 would find it optimal to accommodate, not to fight. So this second Nash equilibrium, it involves a non-credible threat. And if we want to be a little more formal about it, this is going to violate backward induction.

We can easily apply backward induction to this game. We start with player 2. They'd rather accommodate than fight. And given that player 2 is going to accommodate, player 1 would rather enter. And indeed, backward induction gives us the other Nash equilibria of NA.

So this is kind of a review of what we've done before. This is a game of perfect information. We were able to apply backward induction, and that was able to identify one Nash equilibrium as being compelling prediction in the game. And the other Nash equilibrium is being less compelling because it involved a non-credible threat, and more formally, it violated backward induction.

But now let's modify this example slightly. Let's suppose that now player 1 chooses exit or enter as before, 0, 0. But now if player 1 chooses to enter, then player 1 and player 2 will play a simultaneous move game.

And we can represent that by saying, well, player 2 is going to make a move. They're going to choose, say, accommodate or fight. But then player 1 is also making a move. So we'll put these in the same information set. And we'll say player 1 chooses, I think, top or bottom.

So just so we understand what's going on here. Now, as before, player 1 first chooses whether to exit or enter the market. But if they enter the market, then player 1 and player 2 play a simultaneous move game, where player 2 chooses between accommodate and fight, and player 1 chooses between the interpretation is so important, but T and B.

And notice here that I've written this as if player 2 moves before player 1. But whenever we have simultaneous moves, and we want to capture that in the extensive form, the way that we do that is we can arbitrarily pick a sequence and then just assume that player 1 does not observe player 2's move.

So if we had switched the order here, and said, player 1 moves, and then player 2 moves, not having observed what player 1 did, that would be equivalent to player 2 moving first and then player 1 moving not having observed what player 2 did.

So now let me just fill in the payoffs here. Make sure I get this right. So we're going to have 2, 2, negative 1, negative 1. 2, 2 1, 1. Make sure I get this right. And then phi 2t.

Great, so as before, accommodating is good for everyone. Fighting is bad for everyone. But now player 1'a choice between T and B can affect the intensity of that good or bad for the two players.

So now let's go through our same analysis as before. Let's represent this in strategic form, and let's solve for the Nash equilibria of this game. So now player 1 has four strategies, right? Because player 1 chooses N or X and they also choose T or B.

Notice player 1 here only has one information set. If these were in separate information sets, then player 1 would have to move at three information sets. But here we just have one information set and another information set. So for player 1, the strategies are NT, NB-- maybe I'll write it this way, yeah, sure. I'll do it this way. It doesn't matter, XT XB. And the strategies for player 2 are just A and F.

And now we need to fill in the payoffs of this matrix. Well, let's do it systematically. If player 1 exits, then we know that no matter what player 2 or player 1 subsequently does, we're going to get 0, 0. So here we're going to have 0, 0 0,0 0, o and 0, 0. Because in all of these strategies, player 1 first exits, and that's just ends the game with a payoff of 0, 0.

Next, let's figure out what happens under the other two strategies or the other four strategies. So first player 1 enters. If player 2 accommodates, and player 1 plays T then we're going to get 2, 2. And if player 2 accommodates, and player 1 plays B, we're going to get 1, 1. And then if we do it over here, we're going to get-- I'll do that. Yeah, so all I've done is filled in the payoffs in this matrix.

So now let's go through our standard procedure, and let's try to find the Nash equilibria of this game now that we have it in strategic form. We're going to say, if player 1 is going to accommodate, then player 2 finds it optimal to do this. If player 2 is going to fight, then player 1 finds each of these optimal.

And now we're going to go row by row. If player 1 is playing NT, then player 2 would like to accommodate for an NB. Then we're here for here. So now we get three Nash equilibria. Let's write these down. We get NT, A. We get XT, F. We get XB, F.

All right, so here, player 1 enters the market anticipating that player 2 is going to accommodate. And then they follow this up by playing T. In both of these equilibria, player 1 exits the market because they anticipate that player 2 is going to fight. And then it doesn't actually matter. They could do T or B, and we have an equilibrium either way.

But now let's ask the same question as before. Do any of these equilibria seem more compelling to you than any others? Do any of these involve non-credible threats? Any thoughts? Yes?

AUDIENCE: [INAUDIBLE]

IAN BELL: Right, so these two seem to involve non-credible threats. Why? Well, player 2 is fighting when player 1 enters, but fighting just makes both players worse off. So we don't think that if player 1 actually enters, player 2 is going to be willing to follow through and fight.

So intuitively, these involve non-credible threats. But formally, we can't apply backward induction to this game. Why can we not apply backward induction? Yeah?

AUDIENCE: No perfect information.

IAN BELL: It's not a game of perfect information. So the issue here is conceptually, we think that we'd like to select this equilibrium because we think that these equilibria are bad for the same reasons as in our first example up here because they involve a threat that will not actually be optimal for the player to carry out.

However, we don't have the formal methodology to rule out these Nash equilibria because our existing techniques of backward induction don't apply to games of this form because this game does not have perfect information. So the problem here is that Backward Induction, BI, does not apply because this game does not have perfect information. Maybe we'll say, because this game has imperfect information.

What is the imperfect information here? It's that when player 1 moves down here, player 1 has not observed what player 2's action is here. So they don't have perfect information about the previous moves.

And indeed, that makes it difficult to apply backward induction because when we tried to apply backward induction, we wanted to say, well, how should a player move at each of these nodes? But if you're at an information set with multiple nodes in the information set, it's not clear how to carry out this procedure.

So the plan for today, is to introduce a new solution concept of SPNE, which I wrote up here, Subgame Perfect Nash Equilibrium, that is going to capture-- so we'll say this in words. It will capture the idea of credibility, but it will apply to a larger class of games than backward induction did. But it-- to a larger class of games.

Namely, it will apply to games that have imperfect information, and also, it will allow for infinite games, that is, games where we keep making moves over and over again, and there's no clear point.

Remember, we argued last week, when we looked at negotiation, that there was this artificial aspect to our model of negotiation, that there was some time at which the world ended, and there couldn't be any negotiation beyond that point. We had to impose that artificial deadline because we didn't have the tools to analyze games that could potentially go on forever. And now with subgame perfection, we'll be able to do this.

I want to be clear here. Of course, it also applies to games of perfect information and finite games. But in addition to those games, it applies to games of imperfect information and infinite games. So maybe I'll say, also. Of course, it also applies wherever backward induction applies.

And maybe the last thing I'll say, is it extends backward induction. What do I mean by that? Whenever backward induction does apply, subgame perfect Nash equilibrium will give us the same answer. So i.e. it agrees with backward induction whenever backward induction applies.

So we had this class of games where we could apply backward induction. Subgame perfect Nash equilibrium is going to give us the same answers on those nice games, but it's also going to give us a compelling answer in other games beyond that class. And that's what we're hoping to do today.

OK, well, let's try to understand what went wrong, what was non-credible about this. And I'd encourage you one way to think of this-- well, we have the term subgame in the name subgame perfect Nash equilibrium. We're going to think of subgames, that is, smaller games that are part of the larger game.

And notice that once we get to this node, we can think of the rest of this as a game in itself, as kind of a smaller subgame within the larger game. And if we think of this as a smaller subgame, the payoffs in this smaller subgame are actually going to look basically like these.

So we can think of this smaller subgame as, well, player 2 is choosing between A and F, and player 1 is choosing between just T and B. So we can ignore the N. And we can compute what the payoffs are for each of those pairs of actions or strategies in this subgame.

And we can see that only our combination of-- just look at this subgame. So in this equilibrium that we found compelling, in the subgame, player 1 plays T and player 2 plays A. And we can see that in this subgame, if we look at it as a single game, this will actually be a Nash equilibrium of the subgame.

But if we look at how people play in this subgame in any of these strategies, TF or BF, these will not actually be Nash equilibrium of the subgame. And that's going to be kind of conceptually what violates our notion of subgame perfect Nash equilibrium. So let me try to be a little more precise about this.

So what we can do-- so let's apply subgame perfect Nash equilibrium to this example. I said, we can represent this as its own subgame, but let's be a little more formal about this.

We can think of this as we have player 1. They're choosing whether to exit or enter. And if they enter, we now have a subgame which is simultaneous move game, and we can just think of it as a block like this. And indeed, we get 2, 2 negative 2, negative 2.

So this is not technically an extensive form game. This is kind of a heuristic derivation. But we can think of what's going on as after player 1 chooses whether to exit or enter the game, then we play this simultaneous move game. And what we would expect, and what we'd like, is that in this simultaneous move game, the players are actually playing a Nash equilibrium, so we can work within this simultaneous move game and figure out what are the Nash equilibria of this subgame.

Well, we apply our standard procedure. We just underline things. And we see that only this is a Nash equilibrium. So we see that TA is the only one.

So if we believe that once player 1 enters, from thereafter, the two players will play a Nash equilibrium in that subgame, then the only possibility is that thereafter they play TA. And anticipating that they're going to play TA, its optimal for player 1 to enter. And now we've exactly replicated our compelling equilibrium over here. So next I want to describe this procedure more formally. But are there any questions about the example and how we've done this so far in this reduction? No? OK.

So what I argued is that in this big extensive form game, we could treat a subset of this game as a game in itself. And we call that a subgame. And that's what we want to do. So the first step to defining subgame perfect Nash equilibrium, is we first need to define subgames of an extensive form game G.

And the intuition is-- intuitively it's clear. What we want is we want to look at subsets of the tree that would constitute games by themselves if we thought of them separately, all by themselves. So what a subgame is going to consist of, is it's going to consist of a initial-- so a node and all its successors.

So what does that mean? Let's take this as an example. I start here at a node, and I include every node that comes after this. So that's this node, this node, all the way to the terminal nodes, this node, this node, this node, this node.

But we have to have certain conditions. The initial node has to be in its own information set. Maybe I'll emphasize-- so we have a single-- maybe I'll call this the initial node to give it context. So we pick some node, which will be our initial node, and then we include all successors, all nodes that come after that node.

But then we have two requirements. Our first requirement is that the initial node is in a singleton information set. And this is just to ensure that this initial node and everything that comes after it would constitute a game in its own right.

Remember, when we defined an extensive form game, one requirement of extensive form game is that it starts with an initial node, and that initial node is, by itself, in its own information set. So if we want this node to serve as the initial node of our subgame, it better be a singleton information set.

And then the second requirement is that the subgame doesn't break information sets. Let's say what this means in words. What I mean by that is I take a node and all of its successor nodes. I draw a box or a curve around all these nodes, and I need to make sure that there's no information set that includes nodes in this game and also nodes outside the game. That would be what I mean, breaking the information set. And under these conditions, we indeed have a subgame.

So let's note one observation. The entire game is a particular subgame. How can we see this? Why is the entire game necessarily a subgame? What do we choose as our initial node for this subgame?

We just choose the initial node of the game. If we take the initial node of the game and then all its successors, that's just going to be every node in the game. And then we have to check are two conditions satisfied. Well, that initial node better be in a singleton information set because that's part of the definition of an extensive form game. And it can't break information sets because where else could they go? We're including every node in the game.

So let's go back up here and see which nodes can start a subgame. So I'll start here. Can this node start a subgame? Yes, right, this is what we said. If we start here and look at everything after it, this gives us a subgame. And this is precisely the entire game.

And maybe when I say initial node, let me say non-terminal node because I realize we'll get some-- so what about this node? Well, I guess we could think of this as a subgame if we wanted, but there's nothing to do once we're here because we're already at a terminal node, so we're not going to call that a subgame. And that's why we're going to say this is-- we have to start at a non-terminal node.

What about this node? Can this node start a subgame? Yes, right, this is the subgame we already looked at. We started this node. We look at everything that comes after it, and we check that our conditions are satisfied.

Is this initial node in a singleton information set? Yes, and do we break any information sets? Well, the only non-singleton information set we have to worry about breaking is this one, and we see that it's not broken, so we're good.

What about here? Does this constitute a subgame starting at this node? And why not?

AUDIENCE: It breaks infoset.

IAN BELL: It breaks infoset. Actually, it violates both of our conditions. And generally, these can be intertwined. It's going to break an information set, and also, it's not in a singleton information set. So it actually violates both. And then the same is true here. If we tried to start a subgame here, we'd run into problems because this initial node is not in a singleton information set. And we would also break this information set.

OK, so this is how we define subgames. And notice, if you're looking for subgames of a game, because every subgame starts with a non-terminal node, the way you can look for them is just go through each non-terminal node and say, can this non-terminal node start a subgame? So start there.

If we wanted to start a subgame, we're going to have to include every node that comes after it. And then we just check if we do that, do we satisfy these two conditions? So this is how we identify subgames. And now that we understand what a subgame is, we can define what subgame perfect Nash equilibrium is.

So here we'll give our definition. So here we're given an extensive-form game. So let's say we're given an extensive-form game G. Remember, Nash equilibrium could be applied to a strategic-form game. Subgame perfect Nash equilibrium requires us to look at the structure of the extensive form. So we have to be given an extensive-form game here, not a strategic form game.

And we're given an extensive-form game. We say that a strategy profile, S star-- so what we're interested in is a strategy profile NG. Let's remember what a strategy profile is. It specifies a strategy for every player. And remember, in an extensive-form game, what is a strategy? Do you remember what is a strategy in an extensive-form game? Yeah?

AUDIENCE: I believe you define it as a move for every info set?

IAN BELL: Exactly, it's a function or a vector that says for every information set, what action or move do we take at that information set? So I'm not going to include that on the board here. But just keep in mind, that's what a strategy is in an extensive-form game. Is a, maybe I'll just write S, P, and E, so subgame perfect Nash equilibrium if-- well, intuitively what do we want? We want to look at every subgame of our game.

And then we want to look at our strategy profile within that subgame. And we want to check that that would constitute a Nash equilibrium within that subgame. So let's say if for every subgame G prime of G-- so I'm going to impose a requirement on every subgame. And we've already gone through how we can identify all our subgames by looking at all the nodes.

For every subgame G prime of G, well, we want to say that S star is a Nash equilibrium of the subgame. So at first, you might be tempted to say S star is a Nash equilibrium of g prime.

And that's basically what we want to say. But there's an issue here. If S star is supposed to be a Nash equilibrium of G prime, well, the problem is that then that S star must be a strategy profile NG prime. But S is not a strategy profile necessarily NG prime. It's a strategy profile NG.

So S star tells us what we do at information sets that may not be in the subgame. So we just have to be a little more careful and say, well, let's look here. I think we all know what we mean. We want the play specified by this strategy profile to be a Nash equilibrium down here. But of course, we can ignore what's done up here because this is not part of that subgame.

So I'll use the notation, maybe, S star restricted to G prime. And this is just, I don't want to really go through formally the definitions, because I think it just makes it seem harder than it is. But we only look at the parts of the strategies that are relevant inside the subgame. So we might call this the restriction of S star to G prime.

And remember, since we said a strategy specifies a move at every information set, here we're only going to keep track of the moves that information sets that are contained NG prime. Again, this is a graduate game theory class. We get really formal about introducing notation for all these things and what it means, but I think it's pretty clear conceptually what these things mean.

So first question if we have a subgame perfect Nash equilibrium, do we know it's a Nash equilibrium? So if we have a subgame perfect Nash equilibrium, does that imply that something is a Nash equilibrium?

So give me a strategy profile. Suppose I know it's a subgame perfect Nash equilibrium. Do I know that it's necessarily a Nash equilibrium. And maybe a better way, instead of the implication, let's just do this graphically. So what we want to say is it the case that the set of SPNE is a subset of the set of Nash equilibrium?

So we'd like that to be true. Just like when we came over here, we said, there's a lot of Nash equilibria, and we wanted the subgame perfect Nash equilibrium to be a special one of those Nash equilibria. So of course, they could agree, but we want to make sure that every subgame perfect Nash equilibrium, the point is, we want it to be a Nash equilibrium and satisfy these stronger assumptions.

So it should be a Nash equilibrium. How can we see that this is a Nash equilibrium from the definition? So this doesn't require-- you have to see it. But there's not a long argument. It's pretty simple. Yeah?

AUDIENCE: G is a subgame of G.

IAN BELL: Right, so why is this true? Well, G prime G is a subgame. So let's look at our definition. Our definition says for every subgame, this statement holds. So in particular, it holds for the special subgame G prime equals G.

So if we plug that in, it says the restriction of S star to G is a Nash equilibrium of G. But what is the restriction of the strategy profile to G? It's just the strategy profile itself because G is the entire game. So this is stronger than Nash equilibrium, but it implies Nash equilibrium, OK.

Let's maybe go through some examples to see how we would approach this. Let me come down here. Are there any questions before I go to the examples?

So let's do one more example. And this is going to be a variant of our Boston game, where first, we're going to have player 1. And they can choose maybe again, I'll call it entering and exiting. And if they exit, we'll say both players get 1.5.

So we're going to think of exiting. Well, this is just really to illustrate the concept. But this means we're not really going to-- we're going to do some alternative activity, rather than the Red Sox or the Celtics game. But then if they enter, then we're going to play. This is like they say they're sick, and they just stay home.

So now we're going to have, let's say, player 2 is going to move-- choose Celtics or Red Sox. And then player 1, not observing what player 2 does, is going to choose Celtics or Red Sox. And we're going to follow our convention that, I think, player one prefers Celtics to Red Sox, so we're going to have 2, 1 1,2 0, 0. Zero.

So now let's try to compute the SPNE of this game. So here's kind of a recipe for solving SPNE. Step one is always identify the subgames. So how many subgames does this game have? Can we identify the subgames? Well, let's go through.

Starting here is definitely a subgame because the entire game is always a subgame. This doesn't start a subgame because this is a terminal node. So we can't start here. This node will start a subgame because if we go here and everything below it, then we have a single node by itself, and we're not going to break any information sets if we go like that, so that looks good.

What about this node here? Does this start a subgame? No, because this is not a singleton. And then similarly down here, this is not a singleton. So it actually is exactly the same structure as this game over here. So in this case, we're going to have two subgames.

And then a good approach, step two is to work backwards, just like we did with backward induction. So we want to start with the smallest subgames and try to solve for an equilibrium within that subgame and then kind of build up backwards. So step two is going to be, let's work backwards.

So let's see how that works. Well, here, working backwards is pretty easy because there's only two subgames. So we certainly know where to start. If we want to work backwards, we want to start with this subgame.

So let's look at this subgame, and let's represent this in strategic form. So we have Celtics, Red Sox, Celtics, Red Sox. Player 1 is choosing between Celtics and Red Sox. Player 2 is choosing between Celtics and Red Sox. And we know the payoffs are just like the Boston game we studied before.

So this subgame of the larger game reduces to a very simple, simultaneous move game that we can think of in strategic form that we've already analyzed. So what are the Nash equilibria of this game? Let's look at pure. We could also talk about mixed strategies, but let's focus on pure strategies for now. This is the game we've already analyzed, yeah?

AUDIENCE: CC and RR.

IAN BELL: Great, so we have two Nash equilibria. We have CC and RR, right? So notice this is actually a pretty important difference with backward induction applied to perfect information games. When we did backward induction to perfect information games, most of the time, we didn't have ties. We didn't have multiple solutions because at each point, a single player was making a move. And generally one move was better than the other, and we just selected that.

But here, when we're working backwards, we're working at the level of an entire subgame. And we know that generally games can have many equilibria. So when we work backwards, what we're going to do, if we wanted to find all of the Nash equilibria that are subgame perfect, we'd actually need to make a choice here.

We need to first try this and then work backwards and see what happens. And then try this and work backwards and see what happens. And in the end, we may find multiple subgame perfect Nash equilibrium.

So by working backwards, we start at our smallest subgame. We choose a Nash equilibrium of that subgame and then we work backwards. And then we have to do it all over again with a different choice. So you can see, eventually this could get pretty complicated. So let's first make choice one. So here, we have two choices, right? So the first choice would be CC, and the second choice would be RR, in this subgame.

Well now, let's start with our first choice, CC. And we'll do what we did in backward induction is we'll treat this subgame as if once we reach this subgame, this is what happens. So let's highlight CC is going to take us down here.

So now let's move a step back and look at the choice for player 1. So player 1, anticipating that if we get to this subgame, CC is going to be played, and therefore, the payoff is going to be 2, 1, does player 1 want to exit or enter?

AUDIENCE: Enter.

IAN BELL: They want to enter, right? So now when you make choice one, and we work backwards, we see that player 1 wants to enter. And now we have one subgame perfect Nash equilibrium. But we have to be careful to write it down.

Let's see. What is player 1's full strategy? Well, player 1 is going to enter here, and they're going to choose C down here. So remember at the end, we want a strategy profile. So our strategy profile is going to be this.

We have to list how player 1 is going to behave at every contingency. At this contingency they're going to enter. At this contingency, they're going to play C. And actually, I should have highlighted this as well. Even though it doesn't affect payoffs, remember, if player 1 is choosing C, then they have to make the same choice at every node of their information set. So I should have highlighted this. Let me make that clear. So player 1 chooses N here. They choose C at this information set. So we get NC as the strategy for player 1. And then for player 2, we just get C.

Now let's go through the same procedure, making a different choice of our equilibrium in this subgame. So now our choice is RR. I don't want to mess up. I'd have to redraw the graph, so it's going to be too hard to erase. So let's just do it in our heads.

If we choose RR, then upon reaching this subgame, we go RR, The payoff is going to be 1, 2. So now when we take a step back, and we look at player 1, what is optimal for player 1? Yeah?

AUDIENCE: Exit.

IAN BELL: They want to exit because 1.5 is a higher payoff than 1, 2. So now when I go back, I'm going to get exit. And now if I put it all together, what is the complete contingent plan for player 1? It's exit at this node and play R at this point.

Now, so great, in this game, we found two subgame perfect Nash equilibria. But there could be some other Nash equilibria that aren't subgame perfect. Can anyone see an example of a Nash equilibrium that's not going to be subgame perfect? Maybe, it's a bit tricky. But let's just look at this here. Let's focus on this subgame perfect Nash equilibrium.

We know that if player 1 exits, we might say, well, if they're going to exit, it doesn't really matter what they do down here. So you might be tempted to say, let's look at a different strategy profile, say, X C, R. And let's check if this is going to be a Nash equilibrium, or if it's going to be a subgame perfect Nash equilibrium?

So we don't think it's going to be a subgame perfect Nash equilibrium because it didn't get spit out of this procedure. But let's see if it's going to be a Nash equilibrium. Given that player 1 is playing XC as their strategy, can player 2 do any better than choosing R? So does player 2 have a profitable deviation from this strategy? Why, why not? Thoughts here? Yeah?

AUDIENCE: Player 2 doesn't have a choice.

IAN BELL: Right, I mean, they have a complete contingent plan, but if player 1 is playing X, then we're definitely going to end up here whatever player 2 does. So if player 2 were to deviate, the payoff would be exactly the same. And that means that any deviation by player 2, will not be profitable. No deviation is profitable, so player 2 does not have a profitable deviation.

What about player 1? Does player 1 have a profitable deviation? Any thoughts, yeah?

AUDIENCE: Player 2 has to make R [INAUDIBLE].

IAN BELL: And let's see why that is. So we have to be pretty careful here. If player 2 is playing R, well by exiting, I'm getting a payoff of 1.5. Could I ever do better? Well, I could enter and then play C.

So if I enter, player 2 plays R and then I play C, I get 0, 0. That's definitely worse, but I think I have a better deviation. What is my better deviation going to be? Yeah?

AUDIENCE: 3R.

IAN BELL: Yeah, so in fact, instead of XC, I think my best deviation is going to be NR. So instead of doing X, I'm going to enter. And then because I believe that player 2 is going to go to the Red Sox game, I also want to go to the Red Sox game, and that gives me a payoff of 1. But that's still worse than my payoff of 1.5. So indeed, this is a Nash equilibrium, but not a subgame perfect Nash equilibrium.

So if we go to our circle over here, this lies somewhere in here. It's a Nash equilibrium, but it's not a subgame perfect Nash equilibrium. We already checked that it's a Nash equilibrium, or at least verbally. We said neither player can profitably deviate from this strategy profile, and that implies that it's a Nash equilibrium.

We argued it shouldn't be a subgame perfect Nash equilibrium because it didn't come out of our procedure. But let's formally check. What's violated about this strategy profile? Why does this strategy profile not satisfy our definition of subgame perfect Nash equilibrium? So what's wrong? It is a Nash equilibrium, but it's not subgame perfect, and why is that? Thoughts?

Well, let's look at our subgame. In our subgame, player 1 is playing C, and player 2 is playing R in our subgame here. And that means player 1 is playing C. Player 2 is playing R. We're at 0, 0 over here. This is not a Nash equilibrium of the subgame.

Because in this subgame, in fact, either player could deviate. If player 2 changes to R, they would get a higher payoff because then player 1 would get 1. And similarly, player 2 could deviate to C, and they would get a higher payoff.

But we have to be really careful here. I just said that it is a Nash equilibrium and neither player can profitably deviate. And now I'm saying in the subgame, a player can profitably deviate. So what's going on here?

So it is a Nash equilibrium. No one has a profitable deviation in the entire game. But I'm saying, in this subgame down here, there are profitable deviations the players have. And that's why it violates the definition of subgame perfect Nash equilibrium. So what's going on?

So what's going on is there are profitable deviations in subgames that are not reached. So why? Maybe I'll make the more general point. When we're at a point here, what's going on? How can we have a strategy profile that's a Nash equilibrium, but not a subgame perfect Nash equilibrium? What must be the case is that there must be a profitable deviation, and we have to be precise about what that means. Profitable deviation within the subgame that is not reached.

And this is exactly the idea of a non-credible threat. It says, if we were to get to this subgame, if we were to get to this node, some player could do strictly better. That's why this strategy profile violates subgame perfect Nash equilibrium.

But it doesn't violate Nash equilibrium because under this strategy profile, we don't actually get to this subgame. So the fact that I could do better if this subgame were reached, doesn't mean I could do better because when I change my play here, I'm not actually affecting our final payoff because we're going to get stuck over here.

So it's exactly what we talked about with backward induction. Remember, we said Nash equilibrium does not require optimal play at unreached contingencies. Similarly, Nash equilibrium does not necessarily require Nash equilibrium play in unreached subgames. So we're just kind of replacing contingencies with subgames.

And indeed, there may be a profitable deviation at a subgame that is not reached, but that doesn't mean there's a profitable deviation within the entire game. I think that's kind of a subtle point. And that's the heart of subgame perfection. Any questions about that? And if not, we'll move to a more substantive example.

All right, so now let's go over here. And now we're going to consider an application where we can apply subgame perfection to try to make a prediction about behavior in this game.

So this is going to be Cournot competition with entry costs. So we're going to look at classic Cournot competition. Remember, Cournot competition is quantity competition. So each firm chooses how much quantity to bring to market. And then the total quantity brought to market determines the market price, and then each firm sells their quantity at that market price.

And we're going to start with everything exactly as before. So we have n firms. We have constant marginal cost, c. Remember this is the constant marginal cost of production. And then we're going to specify the same convenient demand that we usually specify, which is P of Q.

If quantity Q is brought to market, what is the market price going to be? Well, intuitively, the market price should get smaller the more that's brought to market. And we're going to use the same specification the maximum of 1 minus Q and 0. Exactly what we said before. So if more than one unit is brought to market, the price is just 0. And if nothing is brought to market, the price is 1, and then we interpolate between that linearly.

But now the key difference is that the firms can't just produce immediately. They have to choose whether they want to enter this industry. So I can't just produce corn at my farm today. I have to first buy a farm and start entering the market and decide that I'm going to become a farmer.

So in this example, the firms, each firm has to first make a choice about whether they want to participate in this industry, whether they want to say, pay the upfront costs that would allow them to produce. So in your Econ course, you probably heard the difference between fixed costs and marginal costs. Here, c is playing the role of marginal costs. This is a per unit cost I pay to produce. But I also have something like a fixed cost. I have to buy a factory to start producing or buy a farm to start producing.

| this game is going to proceed in two stages. In stage 1, each firm simultaneously chooses whether to enter the market. In other words, pay the up-front cost to be a player in the market. So each firm simultaneously chooses whether to enter.

And entering has a cost. I'll say, at cost-- well, we'd like to say c, but we already have c here, so I'll say it cost F. You could think of this as a fee or as a fixed cost. This is building a factory, buying the farm.

And then stage 2, the firms who've entered each choose how much to produce. Maybe I'll say, the present firms, that is the firms who've entered, choose quantities. And again, this is simultaneous.

So first, each firm chooses whether to enter at some cost F. Some subset of the firms has entered. And then the present-- this is not great grammar, but I just mean the firms that are present, the entered firms choose simultaneously the quantities to produce. And then we just play the Cournot game.

I would say here, though, we have to be really careful. There's one detail I haven't specified. There's something I need to specify as part of the game. What have I not told you that's important? So we have the players, we have the actions, we have the payoffs, but information.

So we have to be clear here, what do these firms in stage 2 know? Do they know who else has entered? That's an important consideration in the way we model the game. We're going to assume that they do observe this. So I'll say, the set of firms that enter is observed, or maybe I'll just say, entering is observed.

So if you want to start a factory, you want to buy a farm, people observe that. People see that you're now in the market. And that's potentially going to affect the way they behave in the second stage.

And this is crucial. If you were a firm, and you were deciding how much to produce, part of your market research would exactly be to see what other firms are present in the market? And is there a new entrant that may change how much you want to produce? So we're going to assume that this market research has been done, and that this entrance decision is publicly observable.

So now, we'd like to solve for a subgame perfect Nash equilibrium of this game. So let's solve for SPNE. And let's follow our procedure, step 1, identify all the subgames. So any thoughts? What are going to be the subgames of this game?

Well, certainly, the entire game is a subgame. So first, let's find the subgames. Well, we have the entire game. That's always easy. That's what happens at the beginning of stage one. But then there's going to be a lot of subgames in stage 2.

So in stage 2, what's going to start a subgame? A subgame can only start at a history that's observed by the players. That's another way of saying, when we say that the initial node must be in its own information set, think of that as saying a subgame can start at the history if everyone knows we're at that history when we're at it. There can't be any uncertainty about which history we're at.

So what is observed? Well, here, what is the history at stage 2? Well, what we know at stage 2, is which firms have entered and which firms haven't. And in fact, each of those histories will start a different subgame.

So in fact, the subgames are going to be parametrized by the set of firms who entered. I don't have to specify who didn't enter because that's just everyone else. Once I know who entered, I know who didn't enter. And this is in stage 1.

And then we have a subgame parametrized by-- going to say parametrized. That just means we can discard the subgame by this quantity. Maybe I'll call it I subset of 1 through n. And the interpretation of this is the set of players who entered.

And I guess, maybe, if we want to be really formal, let's say it's not an empty subset. Because if no one enters, then we're actually at a terminal node because no one has any decisions to make. So that's not going to start a subgame. So maybe I'll just write it like this.

Sorry, this is math notation. All it means is we have some non-empty subset of firms have entered. We all observe that. And then based upon that, we all decide how to play. Of course, remember, if this is the set of firms who've entered, this is also the set of firms who have a choice to make in the next period. The other firms aren't even present. They don't even matter in this subgame.

So sometimes a subgame might involve only a subset of players. And that's happening here. Only these players are going to be involved in the subgame. Because if you don't enter, you have no choice to make.

OK, so now let's try to work backwards and try to solve this. So step 2 is working backwards. So I've achieved step 1. I've identified the subgames. Now I want to work backwards. So I'm going to start with the subgames here and try to find a Nash equilibrium of each of these subgames.

So suppose we're in stage 2, and suppose that the set of firms that has entered is I. So let's say we're given non-empty subset. So I'm saying, suppose we're at the contingency where this is the set of firms who've entered. That defines a subgame. And now I need to try to solve for a Nash equilibrium of this subgame.

So I need to think, well, what are the substrategies in this subgame? Remember, we just need to think what are the decisions that need to be made within this subgame? Well, what we need is each player firm i in I-- the other firms don't have any decision to make, right? It's only the firms who've entered. Each firm i and I is going to choose a quantity q I to produce.

SO these are the strategies within this subgame. Once I'm in this subgame, each player i who's involved just needs to choose some quantity. And I think we just said this is just some non-negative quantity.

So now we'd like to solve for a Nash equilibrium of the subgame. Do we know what to do here? Does this fit in with something we've studied previously? Any thoughts? Well, this subgame is just the classical Cournot competition game we studied but with a possibly smaller set of firms.

So we already know how to analyze the subgame because the subgame reduces to something we've already studied. So this is just standard Cournot with-- well, effectively, we only have as many firms as have entered. So we need some notation for that. I'll say what this means is, if I have a set I, this means the number of elements of that set. So this is just counting how many firms have entered, so with this many firms.

And we know how to solve that. Notice one thing that all that matters for the Cournot outcome is the number of firms. So even if a different set of firms is entered, each entering firm is going to produce-- the amount that each firm who's entered will produce depends only on the number of firms who have entered, not the identities of the firms.

But that's because of our assumption that the firms are all symmetric. If one firm could produce at a lower marginal cost than another firm, then we'd have to keep track of the identities of the firms. But since all the firms are ex-ante equivalent, ex-ante identical, we don't really care who's entered. We just care how many people have entered.

And what is the solution of this? Does anyone recall what is qi star going to be, anyone remember? This is, I think for tests, a good kind of thing to know. Well, we have 1 minus c on the top because the more costly it is to produce, the less each firm produces.

And on the bottom, well, the more firms there are, the less we produce. We know a monopolist produces 1 minus c over 2. So if we expand that formula, or that pattern, we're going to get the number of firms plus 1.

So if I am the only firm, so the cardinality of I is 1, then I is the only firm and monopolist. And I choose 1 minus c over 2. If there's two firms, we each produce 1 minus c over 3. If there's three firms, we each produce 1 minus c over 4 and so on.

Now let's think back to how we approach this problem before. We looked at the subgame. We found a Nash equilibrium of the subgame. But then the key step was to compute the payoffs in that Nash equilibrium because the payoffs in that Nash equilibrium determined the choices that the players are going to make higher up in the tree. So this is what will actually be done, but let's compute what the payoffs will actually be.

So how do we compute payoffs? Well, my payoff depends on how much I produce and also what the market price is. So let's now compute the market price. So the market price is a function of the total amount produced.

Each of the firms is producing this amount. But there's I of those firms. So the total amount produced is going to be I over I plus 1 times 1 minus c. Do you see we got this, right? Each firm is producing 1 minus c over I plus 1. But there's I firms altogether, so when we add it all up, This. Is the total amount produced.

And now what is the formula for P-- it's actually maybe a cleaner way of doing it is we don't really just care about p. We care about p minus c. Because this is going to be the profit per unit that each firm gets. Because each firm is producing this amount. This is going to be the market price. This is the cost per unit, so the difference is going to be the profit per unit.

And here, we get a nice algebra simplification because we get p is 1 minus q. So we're going to get 1 minus this minus c. But that's 1 minus c minus some fraction of 1 minus c. So we're just going to get the residual fraction of 1 minus c. And we're exactly going to get 1 minus c over this.

You see how this happens? We get all of 1 minus c minus the fraction i over I plus 1 of 1 minus c. So what's 1 minus i over I plus 1? It's just one over I plus 1. If that's tricky, we can go over it again. But any questions on that? No?

OK, and this also makes sense. My profit per unit is decreasing in my cost and also decreasing the number of firms who are entered. So now what we can write down is-- we can actually compute-- so this is a good, I think, thing to remember. What we're computing is the per firm Cournot profit. And it's a bit messy to keep writing cardinality of I, so let's say with k firms.

So what I want to say is if we're playing a Cournot game, and there's k firms all together, what is each firm going to get in equilibrium? Well, it's going to be-- so I'll call this pi k star. And it's going to be the amount produced by each firm times the profit per unit.

So this is going to be-- well, I'll just use this formula --1 minus c over k plus 1. Notice it's exactly the same. I've just changed notation because I don't want to keep writing this. And then I'm going to multiply that by the profit per unit, which is exactly this. Again, I'll use my k.

And it actually is a pretty simple formula that-- I think, this is kind of a good formula to memorize for exams. I mean, you can always rederive it, but I think that's kind of a nice thing to keep in mind.

Let's make sure we see what we're doing here. The first term is the quantity. The second term is the profit per unit. And it turns out that these are exactly the same. Why are they the same? Well, this goes back to you're maximizing area of a rectangle subject to a perimeter, and you always want it to be a square. But that doesn't really matter. This is what you get.

OK, so now we can understand if k firms enter, what's going to happen? If k firms enter, and we play a Nash equilibrium of that subgame, then the k entering firms will each produce this amount. They'll each get this profit per unit, and they'll each get this total profit.

And now we can move to stage 1. So let's go here. So now let's move back to stage 1. So we've actually simultaneously solved. At first, this seemed really hard. We had a huge array of subgames to deal with. But we actually gave one general formula that works for any subgame.

Whatever subgame you give me, I just compute how many firms have entered. That becomes my k. And now I've computed the Nash equilibrium of that subgame. So I've actually covered every subgame at once. And now we need to move to stage 1.

So the question at stage 1 is an entry decision. So now each firm chooses maybe enter N/X, enter or exit. A better term would be enter or not enter. But yeah, enter and exit is cleaner to say.

Well, now, this becomes a bit tricky because let's look at each firm's decision problem. If they exit, that is if they don't enter at all, what is their payoff going to be? Just 0, right?

If they enter, and then after entering, everyone plays a Nash equilibrium, what is their payoff going to be? Well, this is tricky because it depends on how many other firms enter. So we have a formula for their payoff, but it's going to depend on the total number of firms who enter.

So it's going to be pi k star, which equals 1 minus c over k plus 1 squared if the total number of firms who enter is k. Which means if k minus 1 other firms also enter. So if k minus 1 firms also enter, well, when I enter, that makes k, and therefore, this is going to be my payoff.

I think if you just stare at this, it starts getting hard because you've got to figure out the k. The k determines the enter. It gets hard. So I think the best way to do it is to say, can we find an equilibrium where k firms enter? So let's look for an equilibrium.

And if there's an equilibrium where k firms enter, it's not actually going to matter which k. Then there will be an equilibrium for any subset of size k because all the firms are the same. But this may not exist. So let's try to think, if there's an equilibrium in which k firms enter, what must be true?

It means each of the firms that entered, finds it optimal to enter, and each of the firms that doesn't enter, finds it optimal not to enter. So I guess the question is, does any firm have a profitable deviation at stage 1? So let's look for profitable deviations.

Let's consider a firm that entered. What did I use? N. So let's consider this deviation. Let's look at a single firm that did enter. So here k firms, let's assume k is at least 1. Let's look at a firm that entered and consider what would happen if they deviated and exited?

Well, this firm is entering. So what is their payoff from entering? This is their payoff in equilibrium. Their payoff is-- well, what? So let's take the perspective of one of the k firms who has entered. And they're saying, what is my payoff going to be in this game?

Well k firms in total enter, so my payoff is pi k star from entering. But I have to pay this cost, this fee, F. And, I'm sorry. I missed that part, minus F. Because if I enter, I get this in that subgame, but I also have to pay the fee for entering. Otherwise, this would be a pretty easy problem, OK. So this is what I get in equilibrium. What if I exit? Well, then I get 0, right?

So what we need is each of these firms that entered is willing to enter because this is greater than or equal to this. But now you might say, wait a second. If each of these firms found it optimal to enter, doesn't that mean we're going to have a deviation because isn't one of the firms that didn't enter also going to want to enter?

So now it seems like we're stuck. We have to make sure that the firms that entered actually find it optimal to enter. But if they find it optimal to enter, why doesn't one of the other firms that didn't enter deviate and enter as well?

Well, let's go through their problem. Let's look at a firm that didn't enter and consider their deviation. So if they didn't enter, they're getting a payoff of 0. What if one of these firms-- the arrows may not-- hopefully, what I'm saying verbally makes it clear what these arrows mean. You might want to take notes on this.

So if I'm one of the firms that didn't enter, and I deviate now, and I enter, what is my payoff going to be? Yeah?

AUDIENCE: Is it pi k plus 1 star minus F?

IAN BELL: Exactly, and this is the key difference. If I'm one of the firms who's entered, there's only k minus 1 other firms who've entered. But if I'm one of the firms who hadn't entered, now I face a different problem because I face k other firms who've entered. And this means that if I don't enter-- sorry, if I enter, I'll get pi k plus 1 star minus F.

Because there's already k the firms who've entered. If I deviate and enter as well, now there's k plus 1 of us who've entered. And I need this also to be unprofitable. So I need this.

So what is my condition for this equilibrium? Well, F has to be small enough so that the firms that enter find it optimal to do so, but large enough so that the firms that don't enter find it optimal to not enter. And what we need is pi k plus 1 star is less than or equal to F, is less than or equal to pi k star.

And indeed, these pi stars are decreasing in k. Let's look at our formula. This is definitely decreasing in k. So this makes sense. This inequality says, if I'm one of the firms that did enter, and I'm getting pi k star, it was worth it for me to enter because what I'm getting is better than the fee I had to pay to enter.

But this is saying, well, if I'm one of the firms that didn't enter, if I contemplate deviating and entering, I'm only going to get this much, and that can't be worth it for me because F is higher. So let's just draw a little graph to finish.

So what I want to do is, I want to plot all these pi k's on this chart. So we're going to get something like pi 1 star, pi 2 star. It's quadratic, so they're going to start getting closer together pi 3 star, pi 4 star, pi n star.

So we can plot all these points In the graph. And we can try to see what's going to happen based on where F is relative to these points. So given the demand parameters, I can graph each of these numbers all the way up to n because n is the total number of firms.

And let's first say what happens-- we have to be a little careful if F is exactly at a boundary. Let's not worry about that. Let's say what happens if F is here. How many firms enter? Here, no firms enter because this says, the entry cost is higher than monopoly profits.

So the only equilibrium is no one enters. That's an equilibrium because if I were to enter and deviate, then I would get monopoly profits. But that's not enough to offset my fixed cost. So if F is here, we see 0 firms enter. So here, I want to keep track of how many enter.

What if the cost is here? Now how many are going to enter? One, right? This satisfies my equation here with k equals 1. There's an equilibrium with one firm entering. If f is between pi 2 star and pi 1, that's exactly my formula here, so we get that one enter.

And now we can keep going down to here. Two are going to enter. We're going to keep going. And I guess the tricky thing is what happens if F is here? If F is smaller than pi n star, then actually all n are going to enter. And even if F gets all the way down to 0, there's no more firms than n that can enter. So we can't continue this pattern indefinitely because there's only n firms. The most that can happen is they all enter.

So here we've used subgame perfect Nash equilibrium as a solution concept. We've gotten a unique prediction in the game. And we've actually I think, delivered kind of a nice qualitative, industrial organization insight, which says industries that have large fixed costs, will tend to have fewer firms in them. And industries that have very small fixed costs, will tend to have more firms in them.

And this is exactly what we see. We see we tend to see monopolies in industries like cell phone networks and utilities, where you have to pay this huge fixed cost to create your network. But once you've created that fixed cost, the marginal cost is pretty low.

What's going on here? Well, if F is up here, it's worth it for, say, the monopoly utility provider to enter the market. But it's not worth it for anyone else to enter because if they were to enter, they would only-- sorry, I should be here. Let's say fixed costs are here. It was worth it for the first utility company to enter because they're getting monopoly profits that's higher than the fixed cost.

But if a second firm contemplates entering, they're going to have to pay this fixed cost F, but they're only going to get these profits pi 2 star because they're going to have duopoly competition with the other firm, and it won't be worth it for them.

So the simple model delivered, I think, a pretty robust prediction about the organization of industries. I guess one question I'll leave you with, though, is suppose we have n firms, and our equilibrium says k is strictly less than n, should enter. That's our equilibrium. K of them enter. N minus k don't.

Well, which k? How do they coordinate on which k are the mentor? And I think that's kind of a puzzle that maybe this model can't really explain. I'd like to be one of the firms that enters. But if too many of us enter, it's not an equilibrium, so how do we coordinate on this?

Do we mix? Do we flip a coin, decide whether to enter? Is there something missing from the model? Things to think about, but I'll see you everyone on Thursday.