[SQUEAKING]

[RUSTLING]

[CLICKING]

IAN BALL: Today, we're going to start with finitely repeated games. And I want to just start with a really simple example, and we'll actually play a game together. And then we'll develop the general theory of this. So let's recall the standard prisoner's dilemma game where each player had two choices. They could either cooperate with the other player or they could defect. And the payoffs that we wrote down are going to be 2, 2; 3 negative 1; negative 1, 3.

And what we saw in this game is that for both players, defecting is a strictly dominant strategy. So if the opponent cooperates, I'm better off if I defect because 3 is better than 2. And if the opponent defects, I'm still better off if I defect because 0 is greater than negative 1.

But so far, we've assumed that this game is played only once. And in reality, a lot of interactions take place over time. You play a game with someone, you see what happens, then you play the game again. Then you see what happens. Then you play the game again.

And today, we're going to study what happens and how our conclusions change when we're in this kind of repeated situation. So let's start with a really simple example. Let's play the prisoner's dilemma three times.

So how does this work? You're going to simultaneously play the game. Then both players observe the actions that were taken. Then the game is played again. Both players observe the actions that were taken. And then finally, the game is played a third time. And we'll assume that your payoffs in this grand game are just the sum of your payoffs, or maybe the average payoffs across the games.

So if you get 2 in each repetition of the game, your average payoff is 2. I think it's good-- It's fun to just play this. So we could do it on MobLab, but I feel that if you look someone in the eye when you're playing the game, it changes the results a bit. So I want you to pair off, and you need a way to play this.

I think the easiest way is to think of this as like rock, paper, scissors. So just agree, maybe rock-- if you need help-- let's say playing rock is defecting and maybe playing paper is cooperating. So then it's just a variant of rock, paper, scissors.

Play it. Observe the outcome. You play it again, observe the outcome. You play it again. So three times in total. Pair up. Maybe one of you will be by yourself. You can observe other people. And let's do it and see what happens.

All right. So let's see how-- I'm curious how people did. So I hope most groups got a chance to play. So tell me, how did people play? Does anyone want to share what you found when you played this game? Yes.

AUDIENCE: We both cooperated on first try, so I think it has to do with cooperating.

IAN BALL: So you cooperated in all three rounds, or-- OK, interesting. That's pretty good payoffs for you guys. So you each got an average payoff of 2. That's pretty good

I would argue maybe someone was a little irrational there, but we'll see. Anyone else? Did other people cooperate? Yeah?

AUDIENCE: We also cooperated every time, but then after we discussed that because we know each other in real life, like technically the game is infinitely repeated. So maybe that's why we cooperated.

IAN BALL: Great. That's a great point. So we'll see actually that the theory of infinitely repeated games is much different than the theory of finitely repeated games. And Paul Milgrom loves to say the game is always larger than you think. So when you write down these games, often we're missing additional interactions that we might not be modeling.

OK, so a lot of cooperation. Did anyone just defect every time? Yeah?

AUDIENCE: Yeah, I defected all three times.

IAN BALL: OK. Great. A paragon of rationality. And what was your reasoning? Did you have a reason?

AUDIENCE: Well, I thought that Live would also defect all three times because I feel like that's what you're supposed to do.

IAN BALL: OK. Great. So it seems like we have a few different approaches. Some people said, well, we understand that in the staged game, defecting is optimal. So let's just expect that the other player is going to defect and we're both going to defect every period. Other people seem to be able to sustain some cooperation. And maybe what you had in mind is if we cooperate today, then tomorrow, my opponent will reward me by cooperating in the future.

We'll see that in the finite repeated game, the unique subgame perfect Nash equilibrium is actually always defecting, so as one of these groups played. And the way these other groups played is not actually consistent with subgame perfect Nash equilibrium. But let's understand why.

So I think the first key observation here is that when we play a game repeatedly-- so repetition-- I guess I should say, together with observability creates the opportunity for rewards and punishments. I think what some of you might have thought is maybe if I cooperate today, I'll be rewarded, and my opponent will cooperate with me in the future and I'll get some benefit from that. Or conversely, if I defect today, my opponent might punish me by defecting in the future.

Notice that we need both repetition and observability for this to happen. Of course, we need repetition because we need some future for people to reward or punish. And we need observability because the only way to reward and punish people is if you make your future play contingent on what was done in the past. If you couldn't observe the way people played, then the repetition wouldn't really be so important.

Great. So now let's see if we can formally model this game and solve for a subgame perfect Nash equilibrium. The three-copy version actually gets a bit messy, so let me just go over the two-copy version, and then we'll generalize from there. So let's imagine that the game is played twice. So maybe I'll say this is prisoner's dilemma times 2.

And even with this very simple game and just two repetitions, the game tree is going to get pretty complicated, but let's see if we can write it out. So we'll say, maybe the first player is choosing between cooperate and defect. And then the second player, without observing what the first player did, is also going to choose between cooperate and defect.

Now, remember, we've chosen to model it this way as the first player going first, but it doesn't really matter. We could also have the second player going first, and then the first player moving, not having observed what the second player did right. The key thing is that this is really a simultaneous move game, because neither player's action can depend on what the other player is doing.

So this was kind of the first period. Often we call this period 0. We often start our labeling at 0. And now we have the second period. So now player 1 chooses between cooperate and defect.

And notice these are in separate information sets to indicate the fact that. At this point, the players observe what happened in period 0. You know whether both players cooperated, both players defected, one cooperated, the other defected, or vice versa. And then these are going to be in the same information set because this is really a simultaneous move game. And this is player 2. And player 2 chooses cooperate or defect.

You see this gets a bit messy. And I think you see the pattern. And then I could write out all the payoffs, but that's going to be a bit time consuming, so I won't write out the payoffs.

So we want to understand what are the SPNE of this game? And what's the first question we always ask when we try to solve for SPNE? Well, first we have to identify the subgames. That's always step one.

So how many subgames? Maybe I'll write down here. This is period 1. So which nodes of this tree start a new subgame? Well, there's one easy answer. It's this node. I don't know if I have colored chalk today. I'll just use arrows.

So this node starts one subgame that's just the entire game. That's easy. And what other nodes start subgames? Well, we know that if a node is going to start a subgame, it has to be by itself. If a node is with another node in the information set, that node can't start a subgame because that goes against the definition of subgame.

And it has to be a decision node, so it can't be any of the terminal nodes. It can't be any of the nodes that are in multinode information sets. So actually, we immediately see that it's just these four.

So what we see is, first, let's look at the subgames. And it turns out there's five subgames. And here, we see the general pattern. There's one subgame starting in period 0. And then there's four subgame starting in period 1 where each of those subgames corresponds to the history of play that the players have observed.

And in period 1, there's four possible histories-- cooperate, cooperate; cooperate defect; defect, cooperate; and defect, defect. So now let's use our usual approach, and let's try to work backwards. So let's start with our smallest subgames, the subgames that come at the end. And in the subgame starting here, if we're looking for a subgame perfect Nash equilibrium, the restriction of the strategy profile to this subgame must be a Nash equilibrium of this subgame.

So how must we play in this subgame to be consistent with Nash equilibrium? Defect, defect, because we know-- maybe I'll do it here-- that if we fill in this game, we see that DD is the unique Nash equilibrium of this stage game. Now, you might say, well, wait a second. In this subgame, we're not quite playing the stage game, because the payoffs I'm going to put at the end here don't just depend on what happens here. They also depend on what happened in the first game.

But that's not going to change incentives because the payoffs we got from the first game are sunk. We've already got those payoffs. So the only effect that my play has today is on the game that we're playing today. The fact that we add in yesterday's payoffs doesn't change the return to my actions.

So indeed, we must be defect, defect here. And the same reasoning says we have to be defect, defect here, and defect, defect here, and defect, defect here. So what we've seen is that in each of the stage-- maybe I'll say the stage period 1, or-- yeah-- in each of the period 1 subgames, we know that defect, defect must be what's played.

And now let's go to the period 0 subgame. Well, now this is where things are more complicated because here, there's the possibility of rewards and punishments. But I argue that in this period 0 subgame, we still must both play defect, defect in period 0. Why is that? What would happen if we weren't playing defect, defect? What's the key observation here?

Well, I guess the first thing I'd point out is we don't have any rewards or punishments. Why? Because we've already computed what happens in the period 1 subgame. And what we've seen is that both players are going to play defect in period 1 regardless of what was done in period 0. So because the future play is independent of the past play, there's no scope for rewards and punishments. So this is kind of an important point.

Another way of saying this is, how I behave today has no effect on how we're going to behave tomorrow. Because whatever we play today, the same behavior, D, D, is going to be played tomorrow. So how should I play today? If how I play today has no effect on what happens tomorrow, I should simply do today what maximizes my payoff today.

And we know that the unique Nash equilibrium of the prisoner's dilemma is D, D, so in the period 0 subgame we both must play D, D. So that was kind of a bit of a verbal argument, but let's see, any questions on that?

So just to make sure we understand what's going on, what is the strategy-- maybe I'll put it over here. What is the strategy that each player plays? So I might say the outcome of this game is we just always defect. But remember, there's a difference between the outcome and the strategy.

So what is a strategy of a player in this game or in this Nash equilibrium? Well, a strategy is a complete contingent plan. And the number of information sets is actually exactly the same as the histories here. Because when am I called upon to play? Either I know it's period 0 or I'm called upon to play in period 1, and I know what happened in period 0.

So a strategy needs to specify-- we'll write it like this. So we have one spot for each possible information set. Notice I'm being-- if we look really carefully at the game, then the information sets for player 1 are this node, this node, this node, this node, and this node. The information sets for player 2 look a little different. It's this dotted line, this dotted line, this dotted line, this dotted line, this dotted line.

But we can really think of them as being the same, because in both cases, the player either were in period 0 or were in period 1, and all we know is what happened in the first period. So the fact that player 2 comes later and it's dotted versus not, it doesn't actually change things at all. So what is the Nash equilibrium, or the subgame perfect Nash equilibrium we found? It looks like this. This is a full description of the strategy.

So it says initially, before we played it all, I'm going to play defect. Then in the second period-- the first period, the second-- in the next game, if I observe that we both played C, C last game, then I'm going to play defect. If I observe that we played C, D, I'm going to play defect. If we observe that we played D, C, I'm going to play defect, and if D, D, play defect. So this specifies a complete contingent plan in the game.

Notice one kind of subtlety here-- and this will come up. We normally think that how you play in future games will be contingent on how your opponent played in the past. This is how we think about rewards and punishments. If you play C today, I might behave differently tomorrow. And that's possible. But in general, we can actually do even more, that what a strategy allows us to do is make our future play contingent not only on what the other player did, but actually what we ourselves did, which may seem a bit weird, but this will also come up.

I might say, I'll play C tomorrow if I played C in the first period. But I'll play something different if I play D in the first period. That's allowed. We can contingent-- make future play contingent on our own play, and that will actually play an important role below.

So now that we've done an example, let's introduce a general framework for finitely repeated games. You can see this was the simplest possible game. It was prisoner's dilemma. And we only played it twice. And already, the extensive form got really messy. So we're not really going to draw any more extensive forms, and we're going to introduce special notation for repeated games that's a bit more convenient. So maybe I'll call this the general framework for finitely repeated games.

So what's the first ingredient? Well, the first ingredient is the game that we repeat. So this is called the stage game. So in our example, the stage game was the prisoner's dilemma. But it could be any simultaneous move or any strategic form game. And we're going to call the stage game-- we're going to specify it like this. And we might call this game G, which is a strategic form game.

So what does this say? It says in the stage game, player 1 can choose any action from the set A1. Player n can choose any action from the set An. And then we have a payoff function that says, what is player i's utility as a function of the profile of actions that's played in the stage game? And we specify that function for all players I equals 1 through n.

Now, we might call the stage game G. And this is just a strategic form game. You may notice, though, that the notation is a bit different. In the past, earlier in the course, when we wrote a strategic form game down, what was different about the way we wrote it down? We used different notation here. Remember?

These A's are different. Before we wrote Ses, we had S1 up to Sn. And we can think of these as basically the same-- playing the same role as the Ses, but we don't want to get confused between strategies in the stage game and strategies in the grand repeated game. So now I'm going to change notation a bit, and I'm going to call what we do in the stage game an action rather than a strategy. But that's just terminology.

So here, before, we said C and D were strategies of the stage game. Now we're going to say that C and D are actions in the stage game, so that we can keep track of what a strategy is in the full game.

Then we have to specify the timing. And remember this is finitely repeated, so we're going to say that time goes-- we're going to start time at 0. And that will be a bit convenient, we'll see a bit later. And we're going to have period 0, period 1, period 2, all the way up to period T, where T is finite.

So T is sometimes called-- big T is sometimes called the horizon of the game. And notice we actually play the game big T plus 1 times, because we start at 0. And for this reason, we sometimes call the repeated game-- maybe I'll call it G of T.

So the idea is G is our notation for the stage game. G of T denotes the repeated game where we play the stage game G in period 0 all the way up to T. So just so we're on the same page, G of 0 is equal to G, because if we play the repeated game, but a horizon is 0, we just play it once, and that just coincides with the stage game. But in general, this game is going to be more complicated.

So what did we play here? We played g? Of 1. Our horizon was 1 because we played it in period 0 and period 1, and G was the prisoner's dilemma. So as an example, just to remind ourselves what we played so far could maybe be called PD of 1. The game is the prisoner's dilemma, and our horizon was 1.

We next have to say something about information. And what we're going to assume is that past actions are observed. And this was already assumed in our example. What we're saying is whenever we play the game together, we've observed exactly how people played in the past.

So for instance, here, when we played the game for the second time, we observe how people played it the first time. And we pointed out that was crucial. This observability was crucial to create the opportunity for rewards and punishments.

And now we have to say what our payoffs are. Let's use average payoffs. So maybe I'll just write Ui-- maybe I'll be a little more precise. So what I'm saying is, let's say I'm player i, and this is how we play-- or this is the outcome of the game.

In period 0, we play the action profile a0, and all the way into period T, we play the action profile aT. I want to write down what is the payoff of player i? And this is just going to be the sum-- but write it here. It's going to be 1 over T plus 1 of Ui of a0 plus all the way up to Ui of aT

So I look what happens in stage 0, or period 0, I compute my payoff there. I go all the way up to period T, I compute my payoff there. Altogether I have T plus 1 terms, and it's convenient to just divide it by T plus 1. That doesn't have any effect on incentives. We can always scale utilities however we want.

But it means that we're kind of computing utilities in the multiperiod game, in the repeated game, in the same utility units as the stage game. And it turns out that that's convenient, so that our payoff will be something like 2 or 3 rather than 7 or 8, which isn't in the same terms as the stage game.

So now let's talk about strategies. And maybe before we define strategies, let me just make one note that a finitely repeated game, this is a special case of a multistage game. Remember, in a multistage game, we separated the game into stages, and in each stage, some subset of players moved simultaneously having observed what happened in previous stages.

This is a special case because the subset of players who move is always all the players, and the game they play is always the stage game. So more generally, in a multistage game, the stage game could vary over time. We might play a different game in period 0 as we do in period 1. This is a special case in that we always play the same stage game in every period.

So because this is a multistage game, what does that tell us we can apply? Let's just keep in mind, why might this be useful for us? What's a general result we know about multistage games? Remember? If we want to check that something is a subgame perfect Nash equilibrium in a multistage game, it's a bit easier to check because of what result?

IAN BALL: The one-shot deviation principle. So the one-shot deviation principle gave us an easy way to verify that a given strategy profile was, in fact, a subgame perfect Nash equilibrium. The one-shot deviation principle applies to any multistage game. And finally, repeated games are a special case of multistage games, and therefore, the one-shot deviation principle applies to these games as well. And we'll use that fact below.

So now let's define what a strategy is, let's say for player i. And as usual, a strategy is a complete contingent plan. So we have to say how player i will behave at every possible contingency player i finds herself at. And again, let's separate this into periods. We have period 0, period 1, to a generic period t, and then eventually we get to period T.

Well, at period 0, nothing has happened. So we just have the trivial history. So at period 0, my strategy must say Si of the empty set in Ai. So my strategy Si is going to say, what action do I take in period 0?

Well, at the null history we haven't observed anything. Nothing has happened so far. So we're going to use the empty set to denote the history in period 0.

Then in period 1, we need to specify Si of h1. This is going to be an action in Ai. And we have to specify this for every history h1. This is a period 1 history-- h1.

Well, where do these things live? Well, a period 1 history just specifies what happened in period 0. And crucially, not just what I did in period 0, but what my opponents did. So this is going to be an element of A. It's going to be an action profile. Remember, A-- let me write is A1 up to An.

So what this says is in period 1, as player i, I'm going to choose some action to play. And the action I choose can depend upon the way that we played in period 0, namely the action profile that we all took in period 0. If we extend this and we get to period t, what happens? We're choosing Si of ht in Ai.

And again, for every history, but now we have every history, ht. In what? Where does ht live? Well in period t, I observe how everyone behaved in period 0, in period 1, all the way up to period t minus 1. And altogether that's actually t periods. So this is going to be in At.

It might be a little confusing why it's At because it only goes up to T minus 1. But remember, we started at 0. So actions profiles in 0, 1, 2, all the way up to T minus 1. There's t of those altogether. And then we could go down to T.

So maybe if we wanted to say, really, to be really formal here, what is it? A strategy Si is a function that specifies what player i does at every possible history. So maybe we could write this as a function from the union from t equals 0 to T of At to Ai.

So this is, I think, the way to interpret it. But if you want to think formally mathematically, well, I have to specify what I do at every history. So I'm going to take the union of all the sets of histories. And this will specify the period 0 history, all the period 1 histories, all the period 2 histories, all the way up to all the period T histories. And this is going to be the formal way I write it down.

But that's the math way. This is, I think, the intuitive way to do it. And you'll notice, this is maybe the way that we'll write it down in examples that will be a bit clearer. Any questions on the setup on how we've defined a finite repeated game? Yeah?

AUDIENCE: Just a question. So for that game, we're only going for two two rounds?

IAN BALL: Yeah.

AUDIENCE: It just so happened that both players had identical strategies. But normally, we need to specify two of those.

IAN BALL: Exactly right. So I should have said-- yeah, this is a strategy Si. Exactly right. Good point. So the equilibrium that we solved for has to be a strategy profile. So maybe I would say it's actually S1 equals S2 equals this. Yeah. Great. Thanks for clarifying.

That's a symmetric equilibrium. We could also have potentially asymmetric equilibria. Yeah. Great. Any other questions?

So now we want to understand what's kind of special and what happened here. We wrote down this model thinking that there would be maybe a potential for cooperation using rewards and punishments, where people might play differently early on in the game, anticipating what happens later on in the game. But then we looked at the prisoner's dilemma, and we found that didn't happen. We always just played the Nash equilibrium of the prisoner's dilemma every period.

So let's try to understand how general that phenomenon is. And that will be what our first result tells us. So one special property of the prisoner's dilemma is that the stage game only has one Nash equilibrium. It had a unique Nash equilibrium.

And we'll see that when the stage game only has a unique Nash equilibrium, that puts a lot of structure on the subgame perfect Nash equilibria of the repeated game. So let's state this theorem. So let G be a strategic form game. This is going to be our stage game.

So this is an arbitrary strategic form game. May be finite, but I don't think we need finiteness. If this game G has a unique Nash equilibrium, then-- OK. So we've started with an assumption about our stage game. And we've assumed that our stage game G has unique Nash equilibrium.

We want to draw a conclusion about the repeated version of this game. And we actually want a conclusion to hold no matter how many finitely many times we repeat it. So I'm going to say then for every horizon T, the game G of T-- so I'm going to finish my sentence, but let's just make sure we see what we're doing.

We've started with a single game G. We've made an assumption about it, that it has unique Nash equilibrium. And now we're going to draw a conclusion about the T repeated version of this game for every horizon T, namely, the game G of T has a unique subgame perfect Nash equilibrium.

We can actually say more. Any guesses on what this unique subgame perfect Nash equilibrium might be? Yeah.

AUDIENCE: It's G is Nash equilibrium.

IAN BALL: Exactly. And we have to be a little precise, so the set of strategies is a bit different, though, in this game. So let's be a little-- so can you say that a bit more precisely?

AUDIENCE: It's the Nash equilibrium where the action is G is the same as G is Nash equilibrium.

IAN BALL: Yeah. So I would say it's the strategy profile where we play the Nash equilibrium of G at every history. But there's a key point there that we can't just say it's the Nash equilibrium of G, because the set of strategies in G are different. So namely, maybe I'll say-- maybe I'll give notation. Maybe that would be nice.

So unique Nash equilibrium, just add some notation, a star. So I'm just giving a name to the unique Nash equilibrium. This is an action profile. Then this has unique Nash equilibrium, namely S of h equals a star for every history h.

So now we're formalizing what you said. A strategy in the repeated game has to say how we play. A strategy profile needs to say how every player plays at every history. And we're going to specify that whatever the history is, we all follow the Nash equilibrium profile. Maybe if I wanted to be a little more precise, maybe I'll say Si equals ai star for every history h and every player i.

So if I'm player i, what do I do? I have to specify what I do at every history. What am I going to do? Just what I'm supposed to do in the Nash equilibrium of the stage game. And in particular, this strategy is history independent. How I play does not depend on how I played in the past or how other people played in the past. And this means we don't really have any rewards or punishments.

This result is true if the unique Nash equilibrium is mixed, but for the proof, let's think of this as being pure to make it a bit easier, but the result will be true either way. So let's try to prove this. And let's assume that a star is pure, just to make the proof easier, the result goes through, but we haven't really talked about mixed strategies and we don't really want to get into that. So let's just say that this unique Nash equilibrium is pure.

Let me be clear. I'm not saying that it's not enough to just say it's unique among the pure Nash equilibrium. I'm saying there's only one equilibrium, and that equilibrium happens to be pure.

Well, we basically want to replicate the argument that we used in the prisoner's dilemma. So let's suppose that S star-- or S equals S1 through Sn is a subgame perfect Nash equilibrium. So we're going to consider some strategy profile. And let's assume it's a subgame perfect Nash equilibrium, and we want to say that if it is a subgame perfect Nash equilibrium, it has to take this very special form.

Well, what we want to do is we want to work backwards. So let's look at our periods-- 0, 1, 2-- and I'll give a verbal proof. What do we know must be true within period T. How must we all play in the final period, period T? Yeah?

AUDIENCE: You just played [INAUDIBLE]

IAN BALL: Right. And why do we know that that has to be the case?

AUDIENCE: You talked about how all the-- because T is independent of the first t games, so then you just have the one stage played by the Nash equilibrium.

IAN BALL: Exactly right. So once you've gotten to here, it's always easier to work backwards because at the end, there's less to worry about. How I play in the last game can't have any effect in the future because there's no future ahead of me. This is the last game. So my only choice today is, how do I want to influence my flow payoffs in the game today?

And if this is supposed to be-- if this is a Nash equilibrium of every subgame, well, what we know is that we must have S of ht equals ai star for every ht in At. So I'm saying whatever history we're at, in this last period, period T, this history starts a subgame. We have to be playing a Nash equilibrium within this subgame.

But this subgame is just the same as the stage game, because there's only one-- we're in the last period, all we're doing is choosing how to play the stage game once. And therefore, we must be playing a Nash equilibrium. But there's only one Nash equilibrium, so we must be playing that.

And now we want to move backwards. So maybe now I'll go to period T minus 1. So formally, it's kind of a proof by induction, but I'm just going to try to do it verbally. So now what happens in period T minus 1?

Well, in principle, how I behave in period T minus 1 could be more complicated because I have to anticipate that the way I play in period T minus 1 might change the way that we play in period T. The future is more complicated now. But because I know that this is what's happening in period T, I don't have to worry about that. What I know is that if this is how we're playing in period T, then how we're playing in period T is independent of how I play in period T minus 1.

So now I just need to maximize my payoff in the stage game. And the only way that can be the case is if we're playing a Nash equilibrium of the stage game. So if we work backwards we see that Si of hT minus 1 equals ai star for every history ht minus 1 in AT minus 1. And then we can just keep going back until we get to period 0, when we'll see that Si of the empty set must equal ai star.

But notice the ordering really was crucial. It was crucial that we knew what we were going to do in the future in order to figure out our incentives to play today. Yes?

AUDIENCE: The idea that in time T minus 1, because what happens at time T is already fixed. You just play [INAUDIBLE] in and of itself, and then you just look backwards.

IAN BALL: Yeah. So it's exactly backward induction, but I want to be careful when we say fixed. It's not just that it's fixed-- what's crucial is that it's-- yeah-- it's independent of the history that it's-- the crucial thing is that what we do in period t does not depend on hT. If it did depend on hT, then things would be more complicated, because what I do today could affect how we end up playing tomorrow. Because what I do today, I guess the key issue-- and maybe we'll start over here.

So maybe let's go kind of the-- what's the key issue in repeated games? It's that my action ait has two effects. It affects my flow payoff, namely Ui. So if this is how my opponents are playing in period T, the action that I choose directly affects my flow payoff. But it also affects future play.

Why? Well, ait changes the history hs for all S greater than or equal to t because any subsequent history specifies exactly what happened so far. So if I behave differently in period 9, then the history that the players observe in period 10 is going to depend on what I did in period 9. And in particular, that might therefore change actions.

And this is the fundamental trade-off that we always face in repeated games. I have to compare how my action today affects my flow payoffs today, and also how it affects what happens in future play of the game. If we go over here, let's look at what happened in period T minus 1. The first effect was still true-- or this effect is still true. It's still true that at period T minus 1, my action changes the history tomorrow.

But this is what broke down. If the way that we play tomorrow is independent of the history, then the fact that my action changes the history doesn't mean that the actions are going to be different, because the actions tomorrow are independent of the history-- namely this property. So sure, if I change my action at T minus 1, I might change hT. I'm going to change the history that everyone sees tomorrow.

But if the way we play is a star, no matter what that history is, then the fact that I change that history doesn't matter at all. And therefore, I can ignore the effect of my action on future play, and only focus on the effect of my action on my flow payoff. But that means we're effectively just playing the stage game, and we already know that has a Nash equilibrium.

Again, we could do all the algebra, and you can look in the notes for the more formal algebraic argument, but I just find it easier just to think through it verbally than the algebra. Yes, question.

AUDIENCE: Kind of the main point is that if future actions were not dependent on the history, then it might not impact the Nash equilibrium.

IAN BALL: Exactly. And we'll do an example of that exactly next. Yeah. I guess technically what I've shown here-- I should be careful. I've shown that if something is a subgame perfect Nash equilibrium, we must play the stage game Nash every period. I technically also have to check that playing the subgame-- playing the Nash every period is a subgame perfect Nash equilibrium, but that's also not too hard to show.

So I'll just say-- and this is maybe a good exercise. Use the one-shot deviation principle to show that this is a subgame perfect Nash equilibrium. I showed that the only possible subgame perfect Nash equilibrium is this. But of course, it could be that there's no subgame perfect Nash equilibrium. So you should actually check that this is a subgame perfect Nash equilibrium.

And it's exactly the same reasoning. I just have to check that there's no profitable one-shot deviation, and we go through basically the same reasoning. But maybe I'll leave that as an exercise.

So I don't want to give the impression that repetition can't have any effect. And it turns out that this result, it is very, very special that we have a unique Nash equilibrium in the stage game. And if we have multiple Nash equilibria, this result is no longer true. So let's go through that.

So let's go through maybe the oldest game that's been formalized, is we're going to study the stag hunt. So I think Rousseau talked about this in maybe the 1600s. And the story is you have two hunters. They can either hunt hare, which is easy, you do it by yourself, or they can hunt a big stag, and they have to do that together. So can each of them chooses, do we try to cooperate and hunt stag together, or do we just hunt for hare on our own?

And if we both hunt for stag, then we both succeed. We get the stag, and therefore, we each get a payoff of 2. Payoff is going to be 2, 2. Now, hare is not as good as stag. It's just a rabbit.

If we both hunt for rabbits on our own, we each get 1. We each get a rabbit. But a rabbit is worse than a stag. So we each just get 1, 1.

But then the tricky thing is what happens here. What if I hunt for hare as player 1, and player 2 hunts for stag? Well, I'll still get 1 because I can hunt for hare by myself. But now I've kind of screwed over the other person because they're hunting for stag by themselves now. They're not going to be able to successfully get a stag, and they're going to get a payoff of 0.

So the idea is, do we pursue the small prey that we can achieve by ourselves or the big prey that requires cooperation? If you seek out the big prey by yourself, you get nothing and you fail. So this is going to have multiple equilibria. Let's just go through it.

If my opponent is hunting for stag, then I also want to hunt for stag, because if I know they're hunting for stag, then I can join them and we'll get stag, and that's great. But if my opponent is hunting for hare, I also want to hunt for hare, because if I hunt stag by myself, I get nothing. And now we can do the symmetric thing.

And Rousseau didn't really write it out quite this way, but he was trying to make the point that the nature of society can have a big impact. We can be a society of stag hunters, do really well because we all trust each other, or we can be a society of hare hunters who don't do as well. And these are both consistent and stable courses of action. Of course, he didn't have the notion of Nash equilibrium, but he talks about this in his book in the 1600s.

But for our purposes, the key point is that this has two Nash equilibrium, namely S, S and H, H. So now let's consider the repeated version of this. Maybe let's look at stag hunt with T equals 1. So we're just going to play stag hunt twice.

And there's actually going to be a lot of subgame perfect Nash equilibria. Let's look for a few. Let's first look for subgame perfect Nash equilibria where I might say there's no rewards and punishments.

What do I mean by that? I mean let's look for SPNE where the way we play in T equals 1 doesn't depend on how we played in T equals 0. Our play is history independent. Again, if there's no contingency, if the way we play tomorrow doesn't depend on how we played today, then the way we play today, doesn't have any effect on tomorrow, and we could just focus on the stage games.

So any guesses for some kind of simple SPNE that have this property? Maybe I'll write period 0 and period one. There's actually a number of choices here. Any guesses for some really simple SPNE where there's no contingency? We just kind of ignore the past.

Well, one thing we could do is we could just always hunt stag. We know that's a Nash equilibrium of the stage game. So maybe I'll be a little informal, and I'll just write we could do S, S; S, S. So formally, what I mean is in period 0, we each hunt stag. And in period 1, we each hunt stag regardless of the history in period 0. So remember, if I really formally wrote out the strategy here, I'm actually saying S, S at four different contingencies. Great.

So that's one thing we could do. Is this a subgame perfect Nash equilibrium? Well, yes, because in period 1, we're each playing a Nash equilibrium. That's great. In period 0, how I play today has no effect on what happens in period 1. So all we need to check is that I'm playing a Nash equilibrium today. And indeed, that works out. So here's one.

Any other examples of SPNE that don't have any rewards or punishments, that don't have any contingencies? Yeah?

AUDIENCE: We could always hunt hare.

IAN BALL: We could always send hare, right? So again, in each of these subgames, it's obviously a Nash equilibrium. We just have to check that it's a Nash equilibrium of this main subgame. But because the way we play tomorrow doesn't depend on the way we play today, it's Nash equilibrium and we're good.

I'd argue there's actually even some more. Any other things we could do? Yeah?

AUDIENCE: Could you hunt stag in the first period, both of you, and then hunt hares in the next one?

IAN BALL: Exactly. Great. So you might think, wait, oh, this has punishments because we're playing differently tomorrow. But no, punishments are about making what we do tomorrow contingent on what we do today. So what I'm saying, or what our friend here is saying, is we both hunt stag today. Tomorrow we hunt hare, but we hunt hare whatever happened today.

If we both hunted hare today, we hunt hare tomorrow. If we both hunted stag today, we hunt hare tomorrow. If it was S, H or H, S we still do hare tomorrow. So again, there's no contingency and we can go through it. Clearly a Nash equilibrium of this subgame, and then clearly a Nash equilibrium here, because all I care about is my flow payoffs, and S, S is a Nash equilibrium of the stage game.

And then as you pointed out, we have one final one, where we do H, H and S, S. And maybe if I want to be more precise, what I mean-- I mean, at every h1. When I write this down, I'm really specifying four things. I'm saying, at every history h1-- there's four of them-- S, S; H, H; H, S; S, H-- this is what happens.

So here, we have four subgame perfect Nash equilibria that involve no rewards or punishments, and they're pretty easy to analyze. But now we want to see if we can do a little more. Any questions on this before we move on? So now maybe I'll go to this board.

So what's the question? Can we find a subgame perfect Nash equilibrium in which we don't play a stage game Nash in period 0?

So in each of these subgame perfect Nash equilibrium that we looked at, we played a Nash equilibrium of the stage game in period 0. I'm asking, can we find a different subgame perfect Nash equilibrium in which we do not play a Nash equilibrium of the stage game? So let's just try to think through this.

Well, we might say, wait, that's impossible. We can apply our theorem. But why can't we apply our theorem here? Our theorem told us there's only one subgame perfect Nash equilibrium. We already found multiple, so something's going on. So why doesn't the theorem apply to this setting? Yeah.

AUDIENCE: Because it doesn't have a unique Nash equilibrium.

IAN BALL: Exactly right. Our theorem was only applied to games where the-- or stage games with unique Nash equilibria. This stag hunt game has multiple Nash equilibria, so our theorem doesn't apply. Of course, just because the theorem doesn't apply, it doesn't mean that the conclusions aren't true, so we have to be careful. And that's what we're going to do here.

So we're looking for a subgame perfect Nash equilibrium in which we're not playing a Nash equilibrium of the stage game in period 0. So what does that mean? Let's just write out what that means. If we're not playing a Nash equilibrium, then that means some player-- what can they do in period 0? That means some player can deviate in period 0-- unilaterally deviate in period 0 and strictly increase their flow path.

So I realize when I say flow payoff, I just mean the payoff you get in that period. So often we talk about stock and flow. What I mean here is each of these things are flow payoffs, and this is maybe your total stock payoff.

Well, now we have an issue. We're looking for a subgame perfect Nash equilibrium. If it's a subgame perfect Nash equilibrium, in particular, it's a Nash equilibrium, so no one should be able to profitably deviate. But I've already said that some player can deviate in period 0 and strictly increase their flow payoff. So how can we resurrect this as a subgame perfect Nash equilibrium? What has to happen if this player deviates in period 0 and strictly increases their flow payoff?

This can't be profitable, but it increases the flow payoff. So what must happen? Yeah?

AUDIENCE: Maybe in a future one? Would it be a flow payoff is this one, right?

IAN BALL: Yeah, so I'm behaving differently today. I'm increasing my payoff today. But we don't want this to be profitable. So what must happen in the future? I think you're on the right track.

AUDIENCE: It must be profitable then.

IAN BALL: I must be punished in the future. So if the player deviates today, they're increasing their flow payoff, but this is not supposed to be profitable in the entire game. It must be that the deviation they consider today that increased their flow payoff today eventually comes back to bite them, because they're eventually going to be punished in the next period. So what must be true-- so this deviation must be punished tomorrow.

And this is exactly the intuition that I think we all have from just human affairs. If your friend does something that's selfish and hurts you or benefits them today, tomorrow you might punish them by behaving differently. And that's a way of deterring them from engaging in this action. Maybe this is a negative view of human affairs, I don't know. But this I think does happen. Certainly between firms colluding, this happens.

So this is all kind of abstract. Now, what's tricky is we need to punish the deviation tomorrow. But we know that tomorrow is the last period. And we know that in the last period of a game, we have to be playing a Nash equilibrium of the stage game, because this is kind of a key property. In the last period, there's no future. So you must be playing a stage game Nash.

So we need to punish you tomorrow, but we can only punish you by playing a Nash equilibrium. So this seems a little tricky. But the key is that we have multiple Nash equilibrium. So we can punish you tomorrow-- How? by playing a different Nash equilibrium.

And this really gets to the core of the difference between our theorem and our game here. In the last period, we have to play a Nash equilibrium. If there's only one Nash equilibrium, then we know how we have to play in the last period, and there's no scope for punishments or rewards.

If our game has multiple Nash equilibrium like the stag hunt, now we still have to play a Nash equilibrium in the last period, but which Nash equilibrium we play can depend on our past behavior. So there might be a good Nash equilibrium that can serve as a reward, and a bad Nash equilibrium that can serve as a punishment. So let's try to go through that.

So the idea is we want a good Nash equilibrium as a reward. Things could be a bit more complicated, and Nash equilibrium could be good for some people and bad for others. But let's start with this really simple case. And a bad Nash equilibrium as a punishment.

So in this stag hunt game, what is going to be the reward, and what is going to be the punishment? Well, we can look, right? S, S is the good equilibrium, and H, H is the bad equilibrium. So what we want is we want S, S to serve as a reward for good behavior, and we want H, H to serve as our punishment for bad behavior.

So what we say is, if you don't do what you're supposed to today, then tomorrow, we're both going to hunt hare. That's bad for you. If you do do what you're supposed to do, then we're going to hunt stag tomorrow, and that's better for you. So let's see if this works out.

So let's be a bit more specific. I want to find an SPNE in which in period 0 we play something that's not a stage game Nash. Let's be more specific, and let's look for an SPNE where in period 0, we play-- well, we don't want to play a Nash. So that means we're playing one of these. And it's kind of symmetric. it doesn't really matter which one we choose. But let's choose this one. So let's play H, S.

So let's try to fill in our table. I think this is kind of a good way to go about it. So let me-- remember, I think the hardest thing about sometimes finding equilibria is figuring out what kind of object they are, like how many positions, how many things you have to fill in. So let's make a blank table, and then we'll gradually fill it in and try to construct our subgame perfect Nash equilibrium.

So what is our table going to look like? Well, we have our histories. We have the empty history. And then we have four histories after that. We have S, S; S, H; H, S; And H, H. So to be clear, these are period 0 histories, and these are period 1 histories.

And then for each history, we have to specify how player 1 plays and how player 2 plays. So let's put player 1 here and player 2 here. And maybe let's put in a spot. So just to be clear, when I'm asking to solve for an SPNE, I'm asking to specify 10 things, five things for player 1, and five things for player 2, because player 1 has five information sets, and player 2 has five information sets, and we're looking for a strategy profile.

So the goal is to fill this in such a way that we get a subgame perfect Nash equilibrium in which this is what happens in period 0. So we first have a very easy thing to fill in. We want this to happen in period 0. So we're going to fill in H, S. And now H, S is just written this way. So it's player 1 followed by player 2.

So now we need to fill in what happens in period 1 to try to make this a subgame perfect Nash equilibrium. So the first observation is that whatever we do in period 1, at every history, it must be a Nash equilibrium, because we know, in the last period, the subgame is just the same as the stage game because there's nothing else that comes in the future. So what we're choosing, basically, is we need to-- it's almost like kind of a word search.

We have H, H and S, S. These are Nash equilibria. And we need to choose where to put H, H and where to put S, S. So we could be really mechanical about it, but 4-- I mean, then we have 16 different choices, so that's going to take us a while, so let's be a little thoughtful. We could put either S, S here or H, H here, either S, S here or H, H here, and so on.

But that's kind of a silly way to do it. So let's try to think more. Any guesses about where we should put H, H and where we should put S, S. Maybe any thoughts about one position? Hmm?

AUDIENCE: So if he have H, H, we could just-- I mean, if we have S, S, then both people cooperated and we wanted to do so. So like S, S goes under S, S.

IAN BALL: S, S goes under S, S. I'm not so sure. So let's understand. We want a subgame perfect Nash equilibrium where we play H, S in the first period. So you might be thinking we want S, S, but we're trying to get people to play H, S in the first period.

AUDIENCE: Yeah.

IAN BALL: So you said put S, S under S, S?

AUDIENCE: Yes.

IAN BALL: So if we put S, S under S, S, that would make it appealing to play S, S today. So let's think-- so say again?

AUDIENCE: Then H, H would be under S, S.

IAN BALL: Right. Exactly. So let's actually start-- I think this is the kind of-- let's start here. We're trying to encourage people to play H, S today. So we want to reward them if they do play H, S today with a good equilibrium. So let's reward them by playing S, S tomorrow.

And I want to be clear, when we're computing repeated game equilibria, it's very tempting and it's very convenient to think about choosing an action profile rather than just an action. So I'm saying we're going to play S, S tomorrow. But we have to really understand it's always the case that a single player is choosing how they're playing.

We're never jointly choosing what to do. It's just, as the analyst, when I'm analyzing it, it's convenient to think about where we put S, S. But each player is making a decision for themselves. I just want to keep that clear.

So what we've set up so far is we've made it so that if we play H, S, which is how we want people to play. We want this to be a Nash equilibrium. We're going to be rewarded tomorrow by playing S, S. We also need to think about punishments. Let's try to punish the players if anyone unilaterally deviates so.

If player 1 unilaterally deviates, what happens? Player 1 is supposed to play H. If they deviate, they play S And therefore, the history is S, S. So because S, S is a history that could result from a unilateral deviation by player 1, in order to punish that, we want to put H, H here.

Conversely, if player 2 deviates from S to H, then the history is going to be H, H. So we also want to punish H, H with H, H. What about this remaining history? Should we put H, H here or S, S here? Or It doesn't matter.

It turns out it won't matter at all. Why? Well, remember, Nash equilibrium and subgame perfect Nash equilibrium is all about unilateral deviations. So if we're interested in the incentives for people to unilaterally deviate, we have to say what happens if player 1 unilaterally deviates to S, or player 2 unilaterally deviates to H? But there's no unilateral deviation from H, S that will result in S, H.

That would mean that player 1 had deviated from H to S, and player 2 had deviated from S to H. But if they both deviate, well, that's not a unilateral deviation. So it turns out whether this is an equilibrium will not actually depend on what happens here. So we could either do H, H or S, S. Let's do H, H, but S, S would also work.

It can't be S, H, though. It's either H, H or S, S. Or maybe I should have big parentheses to make that clear. So now let's try to check whether this is indeed a Nash equilibrium, a subgame perfect Nash equilibrium.

So maybe I'll call this star. This is our strategy profile. We're not playing a stage game Nash in period 0, but we claim that this is still an SPNE. So I claim that star is a subgame perfect Nash equilibrium.

Now, there's a lot to check. We have to check that this forms a Nash equilibrium in every subgame. So we have to check all five subgames.

But the one-shot deviation principle is going to help us. Technically, we have to check that in every subgame, there's no profitable deviation. But the one-shot deviation principle says we just have to check that there's no profitable one-shot deviation. So check in all five subgames, there is no profitable one-shot deviation.

So the period 1 subgames are pretty easy, but let's just do one of them to check as an example. So in period 1, we have four subgames. Let's look at the subgame following H, H just to make sure we understand.

So let's consider player 1. Player 1 is supposed to play H. At history H, H, player 1 is supposed to play H. So let's compare two things. What happens if player 1 plays H? And what happens if player 1 plays S?

So let's understand what I'm doing here. I'm contemplating a one-shot deviation by player 1 at history H, H. I'm saying suppose we're at history H, H, and player 1 unilaterally deviates. Instead of playing H, they play S, and we want to compare the payoffs.

So what if player 1 plays H as he's supposed to do? What is player 1's average payoff in the entire game going to be? Well, we're at the H, H history. So the way we compute payoffs is we already assume that H, H was played in the first period. So that gives player 1 a payoff of 1.

And then if I play H again, I also get 1 because my opponent is also playing H. So we're going to get a payoff of 1 plus 1 over 2, which equals 1. What if I unilaterally deviate to S? What is my payoff going to be?

Well, if I unilaterally deviate to S-- and maybe let's write-- let me be very clear about what happens here. So if at history H, H, I do what I'm supposed to and I play H, and my opponent also plays H, then this is what our payoff is going to be. The path of play is H, H in period 0, H, H in period 1, and my payoff is an average of 1.

If I unilaterally deviate to S, then after history H, H I unilaterally deviate to S, but my opponent stays at H, which means my payoff is now 1 plus 0 over 2, which equals one half. And indeed, deviating is not profitable. So what I've checked is there's no one-shot profitable deviation for player 1 at H, H. I actually have to do that eight times because there's four histories and two players, so it's a bit of a mess, but it's not too hard to check that that's going to work out.

So now let's go to period 0. This is where I think the action is.

So let's look at period 0. So the history is just nothing. Nothing has happened. And let's look at player 1 and player 2. And in period 0, player 1 is supposed to play H, but they could also play S.

And player 2 is supposed to play S, but they could also play H. So to check that neither player has a profitable one-shot deviation, I need to say, what if player 1, instead of playing H as they're supposed to, one-shot deviates to playing S at this history, and then follows their strategy from there on out. And ask the same about player 2, and check that these deviations are not going to be profitable.

So I think the first thing, before we can compute payoffs, is figure out what happens. So if player 1 does what they're supposed to and plays H today, then what's going to happen is we're going to get H, S today followed by S, S tomorrow. This is what's supposed to happen.

And the same is true down here. If player 2 does what they're supposed to, we get H, S followed by S, S. But what happens if player 1 unilaterally deviates to S, S? Then we get S, S today. But then what happens tomorrow? Yeah?

AUDIENCE: H, H.

IAN BALL: H, H. Exactly. And let's ask the same of player 2. If they deviate from S to H, then we're going to get H, H today. And then what happens tomorrow? Yeah?

AUDIENCE: [INAUDIBLE]

IAN BALL: H, H, right? So now we can exactly see what we were discussing. If player 1 unilaterally deviates, they get a benefit today because S, S is better than H, S. But then they get punished tomorrow because they get H, H tomorrow rather than S, S. And the same thing goes down here.

And instead of adding it all up, I think it's easier to see, what is the gain from the deviation in going from H, S to S, S? Well if they did-- oh, maybe let's just write it all out. Player 1 payoff from H, S is going to be 1, 2; 2, 1. And here, player 2's payoff is 1, ; 0, 2.

So if we add it all up, we're going to get an average of 3/2, 3/2, 1, and 1. So I've just gone through and computed the flow payoff that player 1 gets in each situation. And then I've taken the average. And then I've computed the flow payoff player 2 gets in each situation, and then I've computed the average.

And what we indeed see is that if player 1 unilaterally deviates, well, actually their payoff is going to be exactly unchanged. So it's not profitable. If player 2 unilaterally deviates, their payoff is exactly unchanged. So it's not profitable. And therefore, we see that these deviations are not profitable and we're done here.

And we can exactly see the effect. If player 1 says, wait a second. You're asking me to hunt for hare when my opponent's hunting for stag. I'd much rather hunt for stag and get a higher payoff. Indeed, your flow payoff would jump up by 1. It would jump from 1 to 2 if you deviated.

The problem is that tomorrow, you're going to be punished with H, H, which reduces your payoff by 1. So the punishment you experience tomorrow is just large enough to offset the gain you experienced today. And therefore, the deviation is not profitable. And the same happens for player 2.

Let me stop there, and I'll see everyone on Thursday.