[SQUEAKING]

[RUSTLING]

[CLICKING]

IAN BALL: Great. So today, we're going to talk about something called the folk theorem. So I'll warn you that the class today is a little more in the theoretical side, and then on Thursday, we'll move back to a substantive application. And the folk theorem, in short, says that in infinitely repeated games, we often say anything can happen if the players are sufficiently patient.

Of course, we're going to be more precise about this. When we say anything can happen-- say a little more, what we really mean is any outcome of the game can be sustained as a subgame perfect Nash equilibrium. So maybe I'll say, when I say anything can happen, what I really mean is any outcome can be induced by a subgame perfect-- can be induced by a subgame perfect Nash equilibrium.

An example of this we saw is in the prisoner's dilemma. The outcome cooperate-cooperate could be induced as a subgame perfect Nash equilibrium, but we showed that was only true if the discount factor delta was at least 1/3. And that's going to be a common theme here that anything can happen if the players are patient enough, and that means that the discount factor delta is large enough.

Now of course, not literally anything can happen-- so we'll see, there are some constraints on this, but today, we're going to give a number of theorems of this flavor. And I think you can see how this question arose. Last class, we constructed a special subgame perfect Nash equilibrium that sustained cooperation. But in a more complicated game, you keep constructing subgame perfect Nash equilibria, and the question is, well, where does it stop?

Is there some other outcome that-- could it be constructed, achieved by some subgame perfect Nash equilibrium? Or is there some fundamental barrier to inducing that as a subgame perfect Nash equilibrium? And this theorem helps us address that question. No matter how many equilibria we construct, there's still the question of, are we missing something? Is there something more? And that's what this theorem is getting at.

The name-- I should say a little bit about the name of the theorem. It's not very informative, folk theorem. What it refers to is the fact that this was intuited by game theorists in the 1950s and the 1960s. Everyone thought this was true, but no one had given a formal proof. So it was described as folk wisdom, and today in economics, people often talk about a folk theorem to mean a theorem of this form, but, of course, that really has nothing to do with folk. Folk is just about-- an historical axiom about this result.

And then in the '70s and '80s, this was ultimately formalized-- one of my colleagues here at MIT, Drew Fudenberg, formalized a version of this theorem in 1986 in a very famous paper. So let's start the setting for this. Well, when we say any outcome can be induced or anything can happen, what we're really interested in is payoffs.

So let's start with a simple example. Let's look at the prisoner's dilemma. And let's just recall the basic payoffs we had here. We had CD, CD and we had 2, 2. And it's going to be helpful to visualize these payoffs geometrically. So what I want to do is I want to plot these payoff vectors on a graph. And these payoff vectors have two components, one for each player. So these vectors are going to live in the plane, in Euclidean space of two dimensions.

So let me draw a graph over here. And let's see, we need to go from 3 to negative 1. So let's see how my drawing skills are. Not perfect OK. So let's first plot 2, 2. So that's going to be this point here. Let's plot negative 1, 3. That's going to be here. 3, negative 1, here. And then 0, 0 here.

So these are some payoffs that we can have. And it turns out, it's going to be helpful to connect these. So what we're interested in in the prisoner's dilemma is what payoff vectors can we achieve with some subgame perfect Nash equilibrium? So maybe I'll put the question here. What payoff vectors can be achieved by some subgame perfect Nash equilibrium?

So I think one of them is pretty easy. What about 0, 0? Can we achieve this payoff, 0, 0, in a subgame perfect Nash equilibrium? Well, this one's pretty easy. We know there's an equilibrium where we just always defect every period. And if we always defect every period, that's just repeating the stage game Nash equilibrium. We argued last class, that was a subgame perfect Nash equilibrium, and then we get the payoff 0, 0 every period. So that's something we can achieve.

What about 2, 2? Can this be achieved? So 2, 2 is the payoff if we both cooperate. Well, we showed last class that if we're patient enough, we can find an equilibrium where both players cooperate. But I've been a little, maybe, sloppy here in that, well, why is that called a payoff of 2, 2?

The way we've written it so far, suppose we both cooperate every period. Then what our payoffs are-- technically, what is each player's payoff? Well technically, it's 2 today, plus delta to the 2 tomorrow, plus delta squared times 2 the next day, and so on. And this is not necessarily 2. In fact, it won't be 2. What is this going to be?

Well, we did the algebra last class. We can factor out the 2. And then we have 2 times this series, 1 plus delta plus delta squared and so on, and this is a geometric series and we can compute it, and what we get is 2 over 1 minus delta.

So intuitively, we want to say, if the players always cooperate their payoff is 2 every period, and we want to say that's what their payoff is. But the problem is, the payoff they actually get depends on delta. And we're interested in the case where delta becomes very close to 1, in which case, this payoff is going to get really, really big. But it seems weird to say, well, their average-- their payoff is a lot bigger than 2 because 2 is the highest payoff they can get in the stage game.

And the issue is that we're missing a normalization. What we really want to keep track of is the average payoff. That weighted average payoff, not the sum of the payoffs. So what we're going to do, starting today, moving forward is instead of looking at the sum of the discounted payoffs, we're going to look at the average discounted payoffs.

And the basic issue is, before, what did we do? We computed a payoff stream by saying payoffs are delta-- or, say, for player 1, u1 of a1 plus delta u1 of a2, 0, and so on. So before, what we said is if the action profile in period 0 is a0, the action profile in period 1 is a1, the action profile in period 2 is a2, and so on.

Then we said player 1's discounted payoff is just their payoff in the first period, in the 0th period, plus delta times their payoff in the first period, plus delta squared times their payoff in the second period, and so on.

But the issue is that there's too much weight going on here. And what we now want to look at is what's called the average payoff, which means we're just going to scale this by 1 minus delta. So the average payoff is going to be 1 minus delta, and then let me write this as a summation. So 1 minus delta times the sum from t equals 0 to infinity of delta to the t u1 of a t. And maybe I'll say ui if I want to be more general.

Why this mysterious 1 minus delta? Well, it's to undo the 1 over 1 minus delta that we had before. So the issue is that this is actually a weighted average, but the weights don't sum to 1. 1 plus delta plus delta squared and so on actually sums to 1 over 1 minus delta. So we're just going to normalize things so that things make sense.

And the way we can see if this is a constant-- so let's say this is the same number ui every period, then what we'd like is the average payoff to just equal this number. If I get the same payoff every single period, we'd like my average discounted payoff to equal that number.

And we can see, that's going to work here because-- this chalk is not very good. If we plug this number in here, what do we get? We get 1 minus delta times ui bar times the sum from t equals 0 to infinity of delta to the t. So all I've done is I've said if this is a constant, let me pull it out to the front of the summation.

And now, well, we know that this series is exactly 1 over 1 minus delta, so we exactly get ui bar. So all we've done is multiplied the payoffs we used before by a constant. That's not going to change anyone's preferences. It's not going to change whether anything is actually in equilibrium, but it's going to make it more convenient to keep track of how the players are doing, because we want, if they play cooperate every period, and therefore, get a payoff of 2 every period, we'd like to call that a payoff of 2 in the repeated game, and we'd like that not to depend on the discount factor.

So here, we're going to look at average payoffs. And we indeed see that if the player has played defect every period, then they get 0 every period, their average payoff is 0, and this can be achieved as a subgame perfect Nash equilibrium. We also know that this can be achieved as a subgame perfect equilibrium, cooperate every period if delta is large enough.

What about other payoff vectors in here? Do people have any guesses about what about a payoff vector here? Could I achieve that in a equilibrium? There's no chance. Why? I see people shaking their head.

AUDIENCE: It's not in the outcomes.

IAN BALL: Right. So in each stage game, I'm only getting one of these vectors. So there's no way I can get something up here. And more precisely, what's happening is if in each stage game, in each period, I'm getting one of these, then my average payoff has to be an average of these vectors. And that's exactly what this dotted region is.

So the region in here is exactly the collection of averages of these points. So this may be a little hard to geometrically intuit, but just as a simple example, let's take this point. This point is halfway between 0, 0 and 2, 2. It's the point 1, 1, and I didn't draw that very well, it should be about here. How could I get an average payoff of 1, 1 in this game? Just-- without thinking about equilibrium? Yeah?

AUDIENCE: Just half 0, 0, half 2, 2.

IAN BALL: Yeah. So if roughly, if half the time I played 0, 0 and half the time-- so half the time I played defect, defect and half the time we played cooperate, cooperate then that would average out to 1, 1 We have to be a little careful with discounting because if actually, if I played cooperate here, defect here, cooperate here, it wouldn't quite work out because of discounting, but we could either mix or we could be a little more careful, but it's possible to do this. And it turns out, any point in here can be achieved as some randomization or some weighted average of the different payoff vectors. OK.

So now the question is, which payoff in here could be achieved by some subgame perfect Nash equilibrium? So what we first said is we've narrowed it down to this shaded region. There's no way a vector outside of this shaded region could be achieved by a subgame perfect Nash equilibrium because it's not even feasible. And this shaded region is called the feasible set.

So these are the payoffs that are just-- they're feasible, it's possible to achieve them as some average of these other payoff vectors. But I argue that we probably can't get everything in the feasible set. I think there's some points in the feasible set that I don't see how we could get them as a subgame perfect Nash equilibrium. Any thoughts? What points in here seem unrealistic to you? So we've already gotten this point. We've already gotten this point. What's the point here that we could never sustain as an average payoff of a subgame perfect Nash equilibrium? Yeah?

AUDIENCE: One of the other two points just because why would a player keep on playing for strategy?

IAN BALL: OK. So these points here? I agree. And what's your intuition for this? So let's focus on one of them. Let's focus on this point here-- I have too many arrows. Let me erase these arrows. And let's focus on this one. OK.

AUDIENCE: So player--

IAN BALL: 3, negative 1, yeah.

AUDIENCE: Player 2 would just keep on cooperating because they would just get the worst payoff every time, and they'd start defecting instead.

IAN BALL: Exactly. So this seems really unrealistic. It seems like we can't get this because player 2 is getting a payoff of negative 1. But what if player 2 just always played defect? If player 2 always plays defect, then they always get at least a payoff of 0. So there's no way there could be an equilibrium where player 2 is getting less than 0.

So in fact, we can rule out everything below this horizontal axis, because anything below this horizontal axis is giving player 2 a payoff strictly below 0. But let's look. If player 2 just always plays defect, then, well, they either get 3 if the other player plays cooperate, or they get 0 if the other player plays defect, but whatever happens, they get at least 0, so there's no way they'd be willing to give themselves a payoff this bad. What about the region over here? Yeah?

AUDIENCE: By symmetry, we won't ever have player 1 playing anything that doesn't give them a-- that doesn't give them a non-negative payoff.

IAN BALL: Exactly. So it's the same issue. Here, it's player 1 who's not going to want to do this because player 1 can always play defect every period. And if they do that, sometimes they get 3, sometimes they get 0, but either way, they get at least 0, so why would they ever play in an equilibrium where they're getting less than 0? So we've kind of narrowed things down to this inner region, which maybe I'll shade darker.

So this is a subset of the feasible set. The entire region is the feasible set, but I've thrown out this region and this region to get the darker region here. And it turns out that any point in this darker region can indeed be achieved by some subgame perfect Nash equilibrium if the players are patient enough.

And that's going to be the result that we show. So it's clear that we can't get anything outside this region. That's what I just argued. The hard part is showing that we can get anything inside this region, and that's going to be an implication of the folk theorem, but the folk theorem doesn't apply just to this game, it applies to any abstract game, and that's what we're going to state here.

So any questions? I think this is a little more abstract than some things we've been doing. So if anything about this setup is unclear, this is a good time to ask. Yeah, great.

AUDIENCE: Is there a different definition for the subset of the feasible set that we actually--

IAN BALL: This shaded region? Yeah. So often this is called-- and we'll do a few different versions of-- so let me just preface this by saying the prisoner's dilemma is very special. So if you try to get intuition about this set, you can get misled because it happens to be certain things-- yeah, it's a special case.

But this is often called the set of payoffs that are feasible and individually rational. I don't really like that terminology, but that's often the terminology that's used, yeah. So the whole set is the feasible set, the dark set this is the feasible set together with the constraint that the payoffs are individually rational.

And I should point out, the 0 here, there's nothing magical about 0. 0 only comes up because that's what you get if you defect-- if you-- yeah, if you defect. So it could be shifted a bit. There's nothing magical about 0. It just happens to be that way in this game.

OK. So Now let's try to get to the kinds of results that we're interested in. So let's-- first, let me be a bit more formal about this feasible set. So let's let G be a finite stage game. So it could be the prisoner's dilemma, but it could be something else.

And now I want to formally define what this feasible set is. I just showed it graphically in the example, but let's formally define this. So the feasible payoff set, maybe I'll call it V of G because it's the set of payoff vectors V. Well, what is it? It's the collection of all averages over payoff vectors in here. So what it is it's the set of all sums-- OK, this is a lot of notation, so let's go through it slowly.

What is u of a? u of is a payoff vector that the players get if the action profile a is played. So this is an action profile. And let's just make sure we understand, u of a, what space does that live? So let's say G is a finite stage game with n players. So u of a, what kind of object is this? Yeah?

AUDIENCE: A vector in R n.

IAN BALL: A vector in R n. It's a vector with n components. Because u of a-- remember, this is just u1 of a, all the way up to un of a. So this is a vector that lists what is player 1's payoff. If the action a-- profile a is chosen, all the way up to what is player n's payoff if the action profile a is chosen? So we have an action profile, we have a payoff vector. And now we're taking a summation. We're summing over all action profiles.

So just like over here, maybe we could achieve this point as a weighted average of this point, this point, this point, and this point. That's exactly what we're doing here. We're taking an average over these payoff vectors, we're summing over all the action profiles with possibly some weights. And the weights have to be non-negative and sum to 1. So that's what I mean when I say p is an element of delta A. That just means p is a probability distribution over a. And the feasible set V of G consists of all of these averages.

OK. Great. So we have that notation. And now I think we can-- we're ready to state the form of the folk theorem. I want to remind you of one more piece of notation. Remember that G of delta is our notation for the infinitely repeated game where the stage game is G and the discount factor is delta. So this is the-- with the stage game G and discount factor delta.

So just a reminder here. In general, the feasible set is pretty hard to visualize because it lives in n-dimensional space. But in a two-player game it lives in the plane, and it's a lot easier to visualize. If there are three players, we'd be in normal space, three-dimensional space, and then beyond that, it's kind of tricky.

OK. So let's state-- what I'm going to do here is state-- maybe what I'll call is the folk theorem template. So we're going to state a lot of different versions of the folk theorem, and it's easy to get caught up in the details of these different versions. So I don't want to get distracted by the details. I want us to understand just the general structure of this theorem, and then we'll go through some special cases.

So the folk theorem template is as follows. It says-- here's the statement, let G be a stage game. And now what we're going to do is we're going to pick a feasible payoff vector associated to this stage game. So let's let v be in V of G. So this means that v is a feasible payoff vector in this game G.

And then, here, we're going to say under certain assumptions, which are going to vary. So under certain assumptions, we're going to get the following conclusion. Well, what do we want? We want to achieve v as the outcome of some subgame perfect equilibrium whenever the players are sufficiently patient.

So we're going to say under certain assumptions-- well, what does it mean to say they're sufficiently patient? Well, we just need delta to be large enough. So there's going to be some cutoff delta, and we want delta to be larger than that cutoff. So under certain assumptions, there exists delta bar in 0, 1. So this is going to be our cutoff level of patience. And we're going to be interested in what happens when delta is larger than this cutoff.

So there exists a cutoff, delta bar, such that for all delta greater than delta bar. So, so far, I think it's kind of mathy, but we really haven't done anything. We're just saying, for all-- as long as the players are patient enough, formally there's some number-- maybe it's 0.7, maybe it's 0.3, and we're going to say as long as the discount factor delta is larger than that, what happens?

Well, what we want is we want v to be induced by some subgame perfect Nash equilibrium. So the game G of delta has a subgame perfect Nash equilibrium that yields payoff vector v. So we're going to talk later about what these certain assumptions are, but I just want us to understand, at a high level, what the theorem is saying.

We fixed our stage game, and we fixed our vector that's feasible. So this is some vector that lives in this set over here. And what we want to conclude is that as long as the players are patient enough-- so as long as delta is high enough, if we look at the infinitely repeated game G of delta-- so this is the infinitely repeated game where we play the stage game G every period, and the discount factor we use is delta. In that game, there exists a subgame perfect Nash equilibrium that yields v. What do I mean by that?

Well, in the subgame perfect Nash equilibrium, we can compute what actually happens. There's going to be some sequence of action profiles that are played every period, and we can compute what payoffs every single player gets from that equilibrium. And we want those payoffs to exactly be v.

So let's just-- to make sure we see-- I think this is a bit abstract. We showed a version of this, a special case of this last class. So last class, we looked at the prisoner's dilemma, and we looked at v equals 2, 2. And what we showed is that-- or an implication of what we showed last class is that we could achieve v equals 2, 2 as a subgame perfect Nash equilibrium for delta large enough.

What was the value of delta bar that we established last class? Anyone remember? So we showed a special case of this. For the particular choice of v equals 2, 2, what we did last class implied this. What was the value of delta bar that we could use? Yeah?

AUDIENCE: 1/3.

IAN BALL: 1/3, exactly. What we said is, in the infinitely repeated prisoner's dilemma with discount factor delta, as long as delta was at least 1/3, there was a subgame perfect Nash equilibrium in which the players always cooperated. And when they always cooperated, the payoff vector that they got, the average payoff vector was, indeed, 2, 2.

The folk theorem goes a lot more beyond that because it looks at any feasible vector v, just the particular one, 2, 2, and it applies not just to the prisoner's dilemma, but to any game G. Now, just to make sure we're on the same page here, I've put in brackets here under certain assumptions. We certainly need some assumptions. So why would this theorem-- how do we know this wouldn't be true? What if I covered up these certain assumptions and we just look over here at our discussion before, why would the theorem be false without further assumptions?

Do you remember what we showed? There were some feasible points that couldn't be achieved. We already showed that a point down here and a point up here could not be implemented by any subgame perfect Nash equilibrium. So it must be the case that those points are going to violate whatever these certain assumptions are that we're going to put in here. And then we're going to consider a few different classes of these assumptions.

All right, so let's now try to do a few different versions of the folk theorem. So the way this is kind of developed is there are-- I want people to be able to read this, so let me just leave this here and start a new board. So we want to be able to see this. So let's go here.

So generally, if we make really strong assumptions, then the theorem is going to be easier to prove, but it's not going to get us this far. And the way of the theory has developed is over time, these assumptions have been weakened and weakened to make the theorem stronger and stronger, but then that means the arguments are often more complex. So we're going to start with a really easy version. And this is going to be called the Nash reversion folk theorem. And this one we can prove pretty easily.

So what is the certain assumptions that we're going to need for this version? So the assumptions are-- so these are going to be assumptions about the game G and about this payoff vector v. And our assumption is going to be, there exists a Nash equilibrium-- maybe I'll call it a NE G such that vi is greater than ui and a NE for all players i.

So we know that we're not going to be able to achieve every payoff vector v. We need some restrictions on this payoff vector v. This is one restriction we could have. We say our theorem is going to be true if we have the following assumption. There exists some Nash equilibrium a NE of the stage game G such that every player i gets strictly higher payoff under v than they do under that Nash equilibrium.

So let's try to-- it's a little abstract. Let's apply it here. If we apply this assumption to the prisoner's dilemma, what set will we get? What will be the set of v's that are feasible and also satisfy this assumption? Well, let's go through it.

In the prisoner's dilemma, we know what the Nash equilibrium is. What is the Nash equilibrium of the prisoner's dilemma? DD. They both play D. So in that context, if they both play D, then what is each player get in that Nash equilibrium? 0. So in the context of the prisoner's dilemma, this assumption says we need vi to be strictly greater than 0 for every player i.

So notice that this assumption exactly gives us the darkly shaded region up here because the set of feasible v's that satisfy these assumption are exactly the v's that are strictly greater than 0 here and strictly greater than 0 here, and we exactly get this shaded region. So now we see that this assumption will rule out these bad points over here.

All right, so now, let's see if we can give a proof of the theorem under this assumption. Let me start a new board. So let's give a proof of the natural version, which I'm kind of giving away how we prove it, folk theorem. And proving it in general is a bit tricky, so let's simplify things. Let's look at the special case where v, this payoff vector, is equal to u of a star for some action profile a star.

So in general, v might be only achievable by some lottery over action profiles. And then we have to worry about mixing, and it gets a little tricky. So let's just focus on a payoff vector v that is achieved by some action profile a star, and let's just understand what we want to prove here.

So our assumption is that there's some Nash equilibrium that gives each player i strictly lower payoffs than what they get under this vector v. And what we want to show is that there's some subgame perfect Nash equilibrium that yields this vector v at least if the players are patient enough. So we want to say that as long as the players are patient enough, we can construct a subgame perfect Nash equilibrium of the repeated game that delivers this payoff vector to the players.

So what we need to do is we need to construct an SPNE. Well, we want to construct an SPNE that gives this payoff vector to the players. Well, the easiest way to do that is if we just play a star every period. So what we want is to construct an SPNE that induces the outcome a star every period.

If it induces a star every period, then certainly the players are going to get u of a star every period, and therefore, the average payoff vector will certainly be bv. So then we'll be done if we can show this.

Now this may not work unless delta is large enough, so we're going to have to be a little careful here, but does anyone have any ideas? This is the only thing we know. So how could we use this fact to try to construct our SPNE? Any thoughts here? And maybe the name "Nash reversion" is a bit of a hint.

So let's just think through it intuitively. We want the players to play a star every period. The problem is that a star is not necessarily a Nash equilibrium of the stage game. So it may be that when we're trying to play a star, some player could deviate and strictly increase their payoff today. So we have to have some way to discourage that by punishing that player if that player deviates. Any ideas about how we could punish them? Yeah?

AUDIENCE: Revert to the earlier Nash equilibrium--

IAN BALL: There you go, it's right in the name. Revert to Nash. So what this tells us is that this Nash equilibrium is going to be a punishment. It's a punishment because it gives a strictly lower payoff to every player. So here's our strategy profile. Well, first, we just all play a star until someone deviates. So I'll say each player i plays-- well, a star is an action profile, so player i is going to play their component of that action profile. Is going to play ai star until someone deviates.

So if you remember, the formal way to write down a strategy is to define all the histories, and look at this function s, and it's a real mess. So I'm going to define things informally just using words. And if you have any questions about what these mean formally, you can ask and I'll be happy to answer.

So first, each player i is going to play their component of a star until someone deviates. And then what happens? If someone has deviated, what does player i do? Well, if someone has deviated, we revert to this Nash equilibrium. So each player i plays their component of that stage game nash equilibrium. ai NE.

So let's just see where we are here. I've defined a strategy profile. Let's just check, if we follow this strategy profile, indeed, we're going to get a star every period, because in the first period-- or in the zeroth period, every player is going to play ai star. And then in the next period, since no one's deviated, every player plays ai star, and then the next period, the next period, and so on.

So it's true that if the players follow this strategy profile-- ooh, I missed a y. If the players follow this strategy profile, then indeed, a star is going to be played every period, and therefore, the payoff we're going to get is v. The issue is, this may not be a subgame perfect Nash equilibrium. So we have to check that this strategy profile actually constitutes a subgame perfect Nash equilibrium. And indeed, it's not going to be a subgame perfect Nash equilibrium if the discount factor delta is too low. If the players are too impatient, this is not going to work.

So what we need to show is that as long as the players are patient enough, this strategy profile will indeed constitute a subgame perfect Nash equilibrium, so let's check that. Any questions? I think I-- this may seem a little abstract. OK. So let's go down here.

So we want to check that this is a subgame perfect Nash equilibrium. Or-- in fact, it's only going to be a subgame perfect Nash equilibrium for delta large enough. So we'll check maybe whether-- so remember, whenever you're asked to check whether something is an SPNE, we always follow the same steps. We're going to apply the one-shot deviation principle.

So we're going to check if any player has a profitable one-shot deviation. We have to do that at every single history. But in general, there's way too many histories to look at, so remember, the key step is to group our histories.

And here, we naturally see two groups of histories. There's the histories where no one's deviated and the histories where someone's deviated. So as usual, we're going to group our histories into categories 1 and 2. And it's actually easier to start with 2, so let's start with 2. Someone has deviated. And here, no one has deviated.

So if someone has deviated, how are we supposed to play from here on out under this strategy profile? If someone has deviated, what we do is we just play the stage game Nash equilibrium every single period. And we know that this must be an equilibrium of the subgame because we know that it's always a subgame perfect Nash equilibrium to just play the stage game Nash every period.

So if someone has deviated, what we're doing is we're playing a NE forever after regardless of what people do in the past. Maybe I'll say no matter what. That's key. Why do I say no matter what? if someone has deviated, we're supposed to play the Nash equilibrium from here on out. Once someone is deviated, that can never change. Whatever we do tomorrow, it's still the case that someone has deviated, and therefore, we're going to play the stage game Nash equilibrium.

And we've already argued that this has to be a subgame perfect-- this has to be an equilibrium of the stage game because there's no way you can benefit by deviating from a Nash equilibrium, and because future play is independent of past play. So this is the easy case. This is always going to work.

What if no one has deviated? Well, as usual, let's look at two things. Let's suppose no one has deviated, and let's focus on player i, and let's compare what player I is supposed to do to a one-shot deviation. So as we always say, we're going to say, what happens if they follow the strategy profile or if they choose a one-shot deviation.

So what happens if no one has deviated, and player i follows the strategy profile, and then everyone else follows the strategy profile thereafter? What's going to happen under this strategy profile? Yeah?

AUDIENCE: Payoff is going to be that vector v.

IAN BALL: It is. And let's just first say what the outcome is going to be, and then we can talk about payoffs, yeah.

AUDIENCE: The outcome is that everyone is playing like whatever goes along the vector--

IAN BALL: Yeah, so we had it up here. a star, that's the notation we have, yeah. So we're going to get a star today, a star tomorrow, a star the next day, and so on. Now, what happens if player i chooses a one-shot deviation at this history? Well, before, we were looking at games that only had two actions, so there was only one possible deviation. Now there's actually a lot of different one-shot deviations because player i could deviate to any possible action in the game.

So we have to be more precise. Let's say a one-shot deviation, maybe we'll call this ai prime to show that it's a deviation. So let's look at what happens if at this history, player i deviates to ai prime today, and then follows their strategy profile, their strategy forever after?

So first, let's see what happens today. Well, player i is choosing a unilateral deviation. So today, we're going to have ai prime a negative i star. What does this mean? Everyone else is still playing according to a star. That's what they're supposed to do. Player i is the only one who's deviated, and they've deviated exactly to this action ai prime. That's their one-shot deviation today. There's a lot of different one-shot deviations corresponding to different choices of ai prime, and we're going to have to make sure that none of those deviations is profitable. What happens in the next period, then? Yeah, in the front.

AUDIENCE: The second [INAUDIBLE]

IAN BALL: Yeah. So let's just say, what is the action profile here?

AUDIENCE: [INAUDIBLE]

IAN BALL: Exactly. Now we go to a NE because we're playing Nash reversion. Tomorrow we say, wait, player i is deviated. From here on out, we're going to play the stage game Nash, and we get a NE and dot-dot-dot.

So now we see the classic trade-off. Player i might potentially-- maybe I'll say plus. They might benefit today because by choosing action ai prime, they might strictly increase their stage game payoff, but by doing that, they're going to be punished in the future. I'm going to put a negative here. How is this a punishment? Why do we know that this is worse for player i than this? Yeah?

AUDIENCE: Because they defined over here that vi is strictly--

IAN BALL: Exactly this. This exactly tells us that if we revert to the Nash equilibrium, that will be a punishment for the player. So now-- we'll go through the algebra in a second, but we should just see pretty intuitively, if delta is high enough, and the players are patient enough, the one-shot gain from this deviation is going to be outweighed by the forever after loss that the player-- that the player i is going to experience.

So we can use the threat of punishing players if they deviate to discipline their behavior and prevent them from deviating today from a star, even though a star is not necessarily a Nash equilibrium of the stage game.

Let's actually write this difference out more formally. So let's look at-- maybe-- I'm going to do. So let's write down the gain for player i from a one-shot deviation to ai prime-- it's a deviation, so it can't be equal to ai star, otherwise it wouldn't be a deviation.

So let's define this as delta i ai prime. This says if I'm player i, and I follow this one-shot deviation, how much do I gain? A gain could be negative. So I'm just saying what is gain, but gain could be positive or negative. Well, let's look at it.

Well, we want to write out-- we just want to look at all these payoff differences. So today, my gain is exactly this difference. So today, it's ui of ai prime a negative i star minus ui ai star. So today, I get this instead of this, so my gain is this minus this. Could be positive, could be negative, but we're going to say that's the gain.

And then forever after, I get ui this. Now it turns out that when we're working with average payoffs, we get a really-- this really simple algebra here. So first, I haven't put the coefficients in front. This is what happens today. So if this is what happens today, what needs to be in front of this if I'm interested in average discounted payoff? I don't know if we still have the formula up here, but remember, with average discounted payoff, we-- yeah?

AUDIENCE: 1 minus delta.

IAN BALL: 1 minus delta. Exactly. Because with average, it's 1 minus delta times 1 today, 1 minus delta times delta tomorrow, 1 minus delta times delta squared in the future. And maybe this is a bit harder to see, but if this is what happens forever after, it turns out that this is actually going to be just delta. And one way to see this is that we're working with average payoffs. So if this is how much weight we put on what happens today, then all the rest of the weight has to be on what happens forever after. But if we want to formally see this, let's write it down.

Well, this is-- let's just see how we get this coefficient. So where does this coefficient come from? Well really, it's 1 minus delta times delta plus delta squared plus delta cubed plus our-- right? How do we see this? Because this is the gain I get tomorrow. The gain I get tomorrow is multiplied by 1 minus delta up front, and then delta to discount it for tomorrow.

Then the day after tomorrow is delta squared. Then the day after that is delta cubed, and so on. So if I really write this out, it's this gain I experience times this entire sum. Any questions on where I got this? So maybe to make it parallel, this is 1 minus delta times 1. So what happens today is 1 minus delta times the undiscounted value of 1. What happens in the future is 1 minus delta, but I have to discount it by delta for tomorrow, delta squared, and so on.

But what is this reduced to? Well, this is actually-- let's write it, let's just do a little algebra. This is 1 minus delta times delta times-- well, I'm just factoring out 1 delta from here. Then I get 1 plus delta plus delta squared and so on. But this is just 1 over 1 minus delta. So these cancel, and then I get delta.

So if you follow the geometric series, that's one way of thinking about it. The other way of thinking about it is whenever we're using delta and average discounted payoffs, I put 1 minus delta on today and delta on everything that happens after today. And you can see here, as delta gets close to 1, today is relatively unimportant relative to the infinite future.

Let's be clear, it's always the case that today matters more than tomorrow. I'm not comparing today to tomorrow. I'm comparing today to everything after today. Tomorrow, the day after tomorrow, the day after the day after tomorrow, and so on.

Well, what do we know? We know that this is negative. Why? Well, our assumption over here. We know that what I get from the Nash equilibrium is strictly worse than what I get under a star. So if this is negative, I claim that this is going to be less than or equal to 0 if delta is large enough. Anyone walk me through why-- why this is true? In fact, it's going to be-- maybe to make the argument clearer, I'll say it strictly negative. Yeah?

AUDIENCE: I mean, if you make a really large sum, the first term goes to 0.

IAN BALL: Exactly. That's the argument. Whatever this is, we don't really care what this is, we just make delta really, really big. Then this is basically 0. So this whole term-- first term is basically 0, but then this term is negative, and if delta is close to 1, then we're basically getting all of this negative term, so we're going to get something strictly negative. We could carefully write out the algebra and find the exact bound and be done.

But really, I guess what we've shown is that for each player i, and for each possible deviation ai prime, there's some delta large enough-- let me say maybe delta greater than delta i of ai prime. So really, what this argument is showing is I can choose some threshold-- maybe I'll put a bar-- delta i of ai prime so that if delta is above this threshold, then this particular one-shot deviation for player i is not profitable.

So how do I get my threshold delta bar from the theorem statement? Right now I have a lot of different deltas. I have a different delta for each player i in each deviation, and I'm saying if delta is above that, then this deviation is not profitable. Yeah?

AUDIENCE: Just take the maximum.

IAN BALL: Take the maximum of all of these. So what I know is if delta is bigger than this number, I know this particular deviation is not profitable. But there's another deviation as well, and I have to make sure that's not profitable, but that other deviation also has some threshold delta.

And if I just take the maximum of all these thresholds, then if delta is larger than the maximum of all these thresholds, then it's definitely the case that none of these is profitable. And here, you can see, if you're mathematically inclined, finiteness is playing an important role here because I can only take the maximum over finitely many things. So this is the argument. Just intuitively, take delta big enough, and all of these deviations are going to be unprofitable.

All right, let me-- so that's the proof of the Nash reversion folk theorem. Any questions on that? You look confused, tired. OK. So now let's do-- go a little farther. Maybe I'll call it-- actually, I'm going to save some board space, I'm just going to go straight to here.

So I'm going to look at now what I'll call the individualized nash reversion folk theorem. So before, we just said we're going to punish everyone by going to this Nash equilibrium. But in general, games often have multiple Nash equilibria, and we saw with the Boston game that one Nash equilibrium might be good for one player and a different Nash equilibrium might be good for another player.

So we don't actually need to punish everyone with the same Nash equilibrium. So the idea of the individualized Nash reversion folk theorem is that for each player, we need some Nash equilibrium that we'll use if that player deviates, but that might be a different Nash equilibrium than what we use if another player deviates.

So now, let's say-- let's just rearrange this. Now maybe I'm erasing-- I think now at this point, I might as well just start from scratch. OK. Let's say for each player i, there exists a Nash equilibrium a NE, comma, i of G. So this is an individualized Nash equilibrium.

Such that-- so what does this tell me? It says for every player i, I can find some Nash equilibrium in the stage game-- it could be the same Nash equilibrium for everyone, but it doesn't have to be-- such that that player i does worse under this particular Nash equilibrium than under vi.

So notice, this is a weaker assumption because it may not be that there's a single Nash equilibrium that satisfies this property, but there may be different Nash equilibria for different players. There might be one Nash equilibrium for player 1 that player 1 really doesn't like, and a different Nash equilibrium that player 2 doesn't like.

And the idea is very simple. How do I discourage player one from deviating? I say if player 1 deviates, we're going to play the Nash equilibrium that player 1 doesn't like. And if player 2 deviates, we're going to play the Nash equilibrium that player 2 doesn't like. And that way, neither player wants to deviate. So we're going to use individualized punishments.

So do I have-- let's see. Yeah, great. So here, I think using the old board will actually be helpful. Let's prove the individualized Nash reversion folk theorem. Again, in the special case that v equals u of a star for some action profile a star.

So now, we just need to change step 2. So as before, we start by just playing a star every period, but now the structure of the punishments are different. And maybe I'll say 2i. So if player i deviates, what do we do? Which Nash equilibrium do we play?

Well, if player i deviates, we're going to play this Nash equilibrium that player i doesn't like. So maybe I'll say every player j. So what does this mean? The i here says it's the Nash equilibrium that player i doesn't like. And that itself is an action profile.

So player j is going to play player j's component of the Nash equilibrium that player i doesn't like. And then the same argument is going to go through because when player AI contemplates a unilateral one-shot deviation, they recognize that if they deviate, they're going to get punished with their own personalized Nash equilibrium forever, and if they're patient enough, they're going to be deterred by that.

There's a few issues here, though. We have to be careful. We have to define this a little more carefully. So what happens if a player deviates and then and then another player deviates, what do we do? We have to be kind of careful about that. And it turns out, the trick is-- you only look at who deviated first. Once there's one deviation, we use that punishment forever. So if I want to be more precise, I'll say, instead of if player i deviates, I'll say if there has been a deviation, there's been some deviation, and the first player to deviate was player i.

So if I look back in the history, it could be that there were many, many different deviations. But all I'm going to say is where was the first deviation and who did it? And player i deviated first. Then this is what we play.

So we're getting close. I'd argue there's still a little imprecision here. What could go wrong? What's the one issue with this definition? I said I look back. So how do we play? We look back, we see if anyone's deviated, and then we see who deviated first, and then we punish that person. But what could go wrong here? How do we know who deviated first?

What if two people deviated at the same time? What if, in the 0th period, two players deviate. Now we're in period 1, we look back, we say both players deviate at the same time, who do we punish? Any thoughts? Yeah?

AUDIENCE: Randomly choose which one?

IAN BALL: You could randomly choose, that would be one thing. So if let's say two players deviated simultaneously first, then half the time I'll punish one and half the time I'll punish another. That's a great idea, that's one thing that could work. It turns out, it doesn't even matter what you do as long as you play some Nash equilibrium.

And the reason is, that no one can unilaterally cause two people to deviate. So when players are contemplating whether they should unilaterally deviate, they don't ever take into account what happens if two players deviate at once, but I think your suggestion is the easiest one. So we'll say over here, note, or randomize if two players deviate simultaneously.

So under this strategy, we just each play a star until we see someone's deviated, we look who that person is, we play the Nash equilibrium that's bad for that player, and then we display that Nash equilibrium forever after regardless of what happens again.

Someone else might deviate. Who cares? We just stick with our Nash equilibrium. And because it's a Nash equilibrium, you can check that no one has an incentive to deviate. Or I'll say if more than one, there could be many players [INAUDIBLE]. Any questions about this?

OK. So now we're going to go to the final folk theorem, the deepest folk theorem, I guess. And this is really what was in that paper from 1986. And the question is, can we use harsher punishments than Nash equilibrium?

So, so far, the way we've always punished players for deviating is by reverting to Nash. Either we revert to the same stage game Nash equilibrium regardless of who deviated, or we revert to an individualized Nash equilibrium that depends on the identity of the deviator.

But the question is, can we punish even more harshly than Nash equilibrium? And indeed, there are some games where we want to do that. We talked about in Cournot, the way that we often see punishments is that one player floods the market. One player produces a lot of goods to bring down the market price. But that in itself is not a Nash equilibrium because that player is doing really badly as well.

So the question is, how can we use harsher punishments? And here's the challenge. What was nice is if the punishments were themselves Nash equilibria, then the punishments were self-enforcing. Once we went to that punishment, no one wanted to deviate. But the challenge is, who punishes the punisher for dv?

The problem is, if we want to impose a really harsh punishment-- let's say player 1 wants to really punish player 2, that punishment might be really costly for player 1. But that means player 1 might not want to actually carry out their punishment. So we have to make sure that we punish the punisher if they don't actually carry out the punishment.

But now we get into this infinite regress. Whoever's punishing the punisher, they might not want to carry out that punishment because that might hurt themselves. So we have to have a way to punish the punisher-- the person who punishes the punisher, and then see, we get into difficulty.

So it's hard. What we want to do is we want to say player 1, we don't want you to deviate because if you do, player 2 is going to punish you. But then we have to say, and if player 2 doesn't punish you as she's supposed to, then another player is going to punish player 2. And then we're going to get into these cycles.

So it's going to be quite tricky. And I think for a while, this is why it took people a long time to figure this out, how can we figure out these punishments? And it turns out, the kind of solution, I guess, the simple solution is that we want to do two things at once. Our strategy is going to have two components. The first component is we punish deviations. If someone deviates, the other players punish them. That's easy.

But then, instead of thinking about punishing the punisher for not carrying out the punishment, the trick is actually to reward the punisher for carrying out the punishment. So the general structure is, if anyone deviates, everyone else punishes them. And then if those people actually punish as they're supposed to, then we reward the people for punishing. And this is kind of a high-level idea of how we approach it, and let's go a little more formally.

So to be a little more formal, we want to say, well, what is the harshest punishment, say, on player i? So let's say player I has deviated. We want to punish them. What's the harshest way we can punish player i? Well, intuitively, players, we want to choose a negative i. So a negative i, these are the actions of everyone but player i. We want to choose a negative i to make player i as worse off.

But this is kind of tricky because how bad is this for player i? Well, it depends what player i does. So we have to think, what's the worst way to punish player i? We as players other than player i are going to choose this action profile, but we don't know how bad this is because it depends how player i responds.

But the trick is to say, well, let's suppose player i responds optimally. So let's say we choose a negative i, and then player i chooses the action that's best for him given a negative i. We want to make that as bad as possible. So let's say we choose a negative i, how much can this hurt player i? Well, player i can get the max over ai of ui of ai a negative i.

So this is the punishment. Everyone other than i is trying to punish player i. But player i is saying, well, if this is how you're punishing me, I'm going to choose my action to make my payoff as high as possible. So we can think of this number as representing the strength of the punishment.

So let's think of this as maybe the severity of a negative i as a punishment. The lower this is, the more severe the punishment. So any ideas here? How can we make this punishment as bad as possible?

Well, if this is the severity of this punishment, let's just make this payoff as low as possible. So the trick is to now take the minimum over a negative i. Let's go through this slowly because I think this is a bit tricky.

If we choose a negative i, this is how well player i can do. By minimizing over this, we're choosing the worst possible punishment. We're choosing the action profile so that even when player i plays optimally, their payoff is as low as possible. And we're going to need a word for this. This is going to be vi lower bar. This is basically the worst possible punishment we can impose on player i.

And this is called the pure minmax value. So this is often called minmaxing player i. It means everyone else gets together and chooses the action that makes player i's payoff as low as possible assuming that player i best responds to this punishment. So now we can state the final version of the folk theorem. Maybe I'll put it here, and then we'll conclude.

So this is maybe called the pure minmax folk theorem. And in the Nash reversion folk theorem, we just had to make sure we could punish player i by reverting to the Nash equilibrium. Here, we're going to make sure we can punish player i by using the worst possible punishment. So the condition, the assumption is simply going to be that vi is greater than vi lower bar for all players i.

And it turns out that this assumption is weaker than all of the other assumptions, and as a result, this version of the minmax of the folk theorem is stronger. There's a larger range of v's that satisfy this condition. Because vi may not, in this case, be higher than any Nash equilibrium payoff, but it is higher than this worst possible punishment over here.

It turns out, we need one technical condition, so I'll write this in parentheses. And V of G-- remember, V of G is the set of feasible payoff vectors, has full rank. This is a technical assumption that-- it's fine. It's not covered-- it's not going to be tested, I just don't want to write something wrong and have people complaining, so there is one assumption.

And this comes from the paper from 1986, so now we're really getting to the frontier. And let me just give you a really simple outline of the proof following what we described. So the proof sketch, now there's going to be three stages. So initially, everyone's going to play-- maybe I'll just write it a bit more simply. We play a star. Remember, a star, as usual, is going to be the action profile that gives us the payoff vector v. We're going to play a star until a deviation.

And then we have to specify what happens if player i deviates. And it turns out, there's going to be two components of this. You might say, oh, if player i deviates, then we're just going to impose the worst possible punishment on player i forever. But the problem is, people might not want to carry out that punishment because it may be very costly for them to punish. Remember, we have to also reward the punishers.

So the trick is we kind of minmax Player i for n periods. And then, if we get through this stage, we reward everyone for actually carrying out the punishment. Forever after. But the structure of this is a little more complicated.

So first, we play a star. Then we're going to minmax, meaning impose the harshest possible punishment on player i for n periods. And then, if we actually carry that out, we reward the punishers forever after. But if at any stage any player deviates, we go back to the beginning of stage IIi. If player j deviates at any point, now we minmax player j for n periods and then reward the punishers. If any of the punishers deviate, we go back to the beginning of that for that corresponding punisher.

So if at any stage-- and I'm not going to write this on the board, but I'll just say it in words. If at any stage of this strategy profile a player deviates, we start this sequence for that player. We minmax them for n periods, and then we reward the players for carrying out that punishment. And if someone else deviates, we go back and do it again.

And this is a much more complicated structure. People who are still professors at MIT came up with this, and this is getting to the frontier here, and this is much harder to deal with.

But let me-- so I'll say, I wouldn't expect you to be able to fully understand this proof on an exam, but we may ask-- I may ask a little bit about it on a problem set. So let me stop there, and I will see everyone on Thursday.