

本節和下節(即 RL4)的總結:

2-player, zero-sum, deterministic game of perfect information: $\text{minmax} = \text{maxmin}$

non-deterministic: $\text{minmax} = \text{maxin}$, nothing different.

non-deterministic, hidden information: $\text{minmax} \neq \text{maxmin}$

(mixed strategy: choose the peak of the Λ shaped curve)

non-zero-sum, non-deterministic, hidden information: Nash

above number of rounds known: Nash

above number of rounds unknown: Nash + subgame perfect equilibrium (independent of the history, plausible threat)

zero-sum stochastic games: in Bellman Equation, replace max with minimax, 'value iteration works' 等幾個性質都成立

general stochastic games: in Bellman Equation, replace max with Nash, 'value iteration works' 等幾個性質都不成立

GAME THEORY

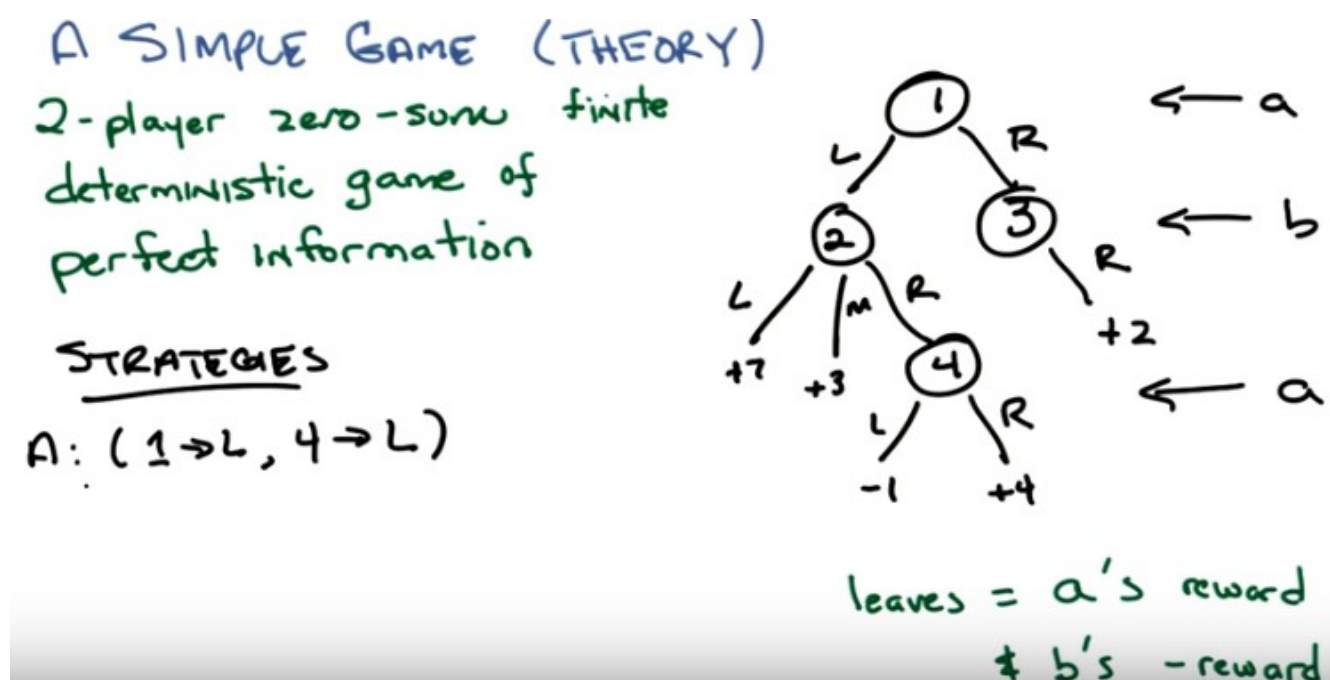
1. Hi Michael. >> Hey Charles. How you been? >> I've been doing just fine. How've you been? >> Good. It's been a while. I understand you went to Africa. >> I went to Africa. The entire continent, or at least parts of it. >> How's Africa? >> It's fine. It's, warm. >> The entire continent? >> [LAUGH] The entire continent. Or at least Kenya and Nigeria. >> Okay. >> Particularly in Nigeria. Which was 80 something degrees and 4,000% humidity. [LAUGH] >> So, do you want to do a shout out to any of, any of your friends from there? >> Yeah, I'd like to say hi to Chairman and I'd like to say hi to Prof. Oh, and Good Times. I would like to say hi to Good Times. He was awesome. Okay, so Michael. I am back now so that we can talk about the last little section. That we're going to do in this class, and it's game theory. >> That sounds cool. What does that have to do with machine learning? >> So that's an interesting question because, in fact, game theory comes from a tradition way outside of machine learning in A.I. but you'll see in a moment why it is I care about game theory, and really this is a very natural extension to all the stuff we've been talking about with reinforcement learning. >> Interesting. Are we talking about games like Monopoly? >> In some sense, we are. Because all of life is like Monopoly. So in fact, what is game theory? Maybe that's what we can do. We can just try to define game theory, and maybe it'll be clear why it is we're worried about this at all. Seem fair? >> Yeah. >> Okay.

GAME THEORY

- MATHEMATICS OF CONFLICT
- SINGLE AGENTS →
MULTIPLE AGENTS
- ECONOMICS (& POLITICS ...) ^{SOCIOLOGY} _{BIOLOGY}
- INCREASINGLY A PART OF AI/ML

2. Alright Michael, so there's lots of definitions of game theory that we could use. One that I like in particular is that game theory is the mathematics of conflict. >> Hm, [CROSSTALK] that's interesting. >> I think it's kind of interesting. Or generally it's the mathematics of conflicts of interest when trying to make optimal choices. >> because I feel like a lot of people have their own conflicts with mathematics. >> I think everyone but mathematicians have their conflicts with mathematics. I think that's fair. >> I see. >> But do you see if you, can you see how worrying about the mathematical conflict might be a sort of natural next thing to think about after you've learned a lot about reinforcement learning? I guess then well the next bullet kind of, kind of suggests a trend. So, so we've been talking about decision making and it's almost always in the context of a single agent that lives in a world and it's trying to maximize reward. But that's kind of a lonely way to think about things, so what if there's other agents in the world with you? >> Right and of course evidence suggests that there are in fact other agents in the world with you. And what we've been doing with reinforcement learning which, you know, has worked out very well for us, is we've been mostly pretending that those other agents are just a part of the environment. Somehow all the stuff that the other agents do is hidden inside of the transition model. But truthfully it probably makes sense if you want to make optimal decisions to try to take into account explicitly the desires and the goals of all the other agents in the world with you. Does that seem fair? >> Yeah. >> Right. So that's what game theory helps us to do and then at the very end I think we'll, we'll be able to tie what we're going to learn Directly back into the reinforcement learning that we've done and even into the Bellman equation. >> Oh, okay, nice. >> Yeah, so that is going to work out pretty well but, but we have to get there first and there's a lot of stuff that we have to do to get there. But right now what I want you to think about is this notion that, [we're going to move from reinforcement learning world of single agents to a game theory world of multiple agents](#) and tie it all back back together. It's a sort of general note that I think that, that's worthwhile is that, game theory sort of comes out of economics. And then in fact, if you think about multiple agents there being millions and millions of multiple agents, in some sense that's economics. Right? Economics is kind of the math, and the science, and the art of thinking about what happens when there are, lots, and lots, and lots, and lots of people with their own goals possibility conflicting, trying to work together to accomplish

something, right. And so what game theory does, is it gives us mathematical tools to think about that. >> I feel like I feel like other fields would care about some of these things too, like sociology. >> Right. >> And what about, I could kind of imagine biology caring about these things, too. >> Even biology, I like the idea of biology. Biology. Why would biology care about this? >> Well, I guess the way you described it in terms of lots of individual agents that are interacting. Like, you know, creatures that live and reproduce. I feel like they, they have some of those same issues. >> Sure. So certainly biology at the level of entities, at the level of mammals or level of insects, you might be able to think about it that way. But perhaps even at the level of genes and at the level of cells. Little virii and, and bacteria. You could possibly think about it that way. >> because they're in conflict too, I guess. >> Yeah. Now there's probably this notion of intention. It's not entirely clear what that means here and I think implicit in the notion of what we're doing here is this notion of intention and explicit goals as opposed to ones that are kind of built into your genes, but I think that's a perfectly reasonable way of thinking about it. I think the, the lesson from this discussion though is that. What game theory sort of captures for us or what would like for it to capture for us, is ways of thinking about what happens when you're not the only thing with intention in the world and how do you incorporate other goals from other people who might not have your best interest at heart or might have your best interest at heart. How do you make that work? And so if you think about that problem then I think it makes sense that it's been increasingly a part of AI over the years, and in some ways machine learning has started to think of it as being a mainstream part of what we do. >> Cool. >> Hence, why it's worth talking about today. Okay. Sound good. >> Gotcha. >> Okay. Let's, let's try to make this concrete with a very simple sort of example.



上圖右邊的 game 意思沒說清楚(幾乎沒說), 害得老子看了好久才看懂。

意思就是: 此 game 中有個球, 它從節點① 出發。有兩個玩家(agents): A 和 B, 它們輪流控制球怎麼走。例如球最開始在①, 玩家 A 讓球往左走到②, 然後玩家 B 讓球往右走到④, 然後玩家 A 再讓球往右走。

由此可知,

當球在第一層(① 那層)時, 它由 A 控制;
當球在第二層(② 和 ③那層)時, 它由 B 控制;
當球在第三層(④ 那層)時, 它由 A 控制;
不管誰讓球走到 leaf, 都是 A 得到該 leaf 的值, B 得到該 leaf 值的相反數。

3. Okay, Michael, so here's a, a nice little concrete example to, to think about this. Let's pretend that we're no longer in a world of a single agent like we've been thinking about with reinforcement learning, but we've gone full-blown generality to two agents, [LAUGH] okay? And let's call those agents a and b, and they're going to be in a very simple game where A gets to make a choice. And then B gets to make a choice. And then A might be able to make a choice. So this tree that I've drawn over on the right is going to capture the dynamics of this specific game that I'm imagining. So, these little nodes here, these circles represent states. And we can think about those in states in the same way that we. Talked about reinforcement learning in the past. The edges between these nodes represent actions that one could take. So, this should look familiar, this is just basically a game tree like anyone who's taken a, an AI course might have seen. Okay? >> I guess so. It doesn't look like a very interesting game. >> No. >> But I guess it's a, sort of abstract example. >> Yes. It's a very simple game just so that we can get a handle on some basic concepts. So, in particular, if you look at the details of this game, you start out in state one. Ok? And A gets to make a choice between two actions, going left or going right. If A goes right, goes right, she ends up in state three. If she goes left, she ends up in state two. Regardless B gets to make a choice. From state three we can choose to go right, and really that's all that can happen. And this, what happens if B goes right from state 3 is that, a value of plus two is assigned to A, okay? All of these numbers at the bottom, all the leaves here, are going to be values or rewards if you want to think about 'em that way that are assigned to player A. And, in fact, for the purposes of this game, it's going to be the case that B always get's the opposite of what A get's. So, if A get's +2 then B get's -2, if A get's plus 4 then B get's minus 4, if A get's minus 1, B get's plus 1, does that make sense? >> Yeah, though could you write it down so that I won't forget? >> Okay, that's fine. So, by the way, this is a very specific type of game. here, and it has a name, which I want to get right. This is a two-player zero-sum finite deterministic game of perfect information. So as a, as a title or description of, of this kind of game, does this make sense to you? Do you think you know what they all mean, what all those words mean? >> So, two players because it's a and b, zero-sum. Because you said the leaves are a's rewards and b's reward is the negation so if you add the two rewards together you're always going to get zero. >> That's almost right. >> [LAUGH] Ok. >> It's not exactly right. Actually, so zero sum really just means that the sum of the rewards is always a constant. >> And that constant needs to be zero. >> It doesn't need to be zero. >> So if it added up to 11, that would still be zero sum? >> If it added up to 11 everywhere. Yes. >> Huh, okay, interesting choice of terminology. finite, I don't know, everything seems to be finite here. There's no infinite number of choices or states or depth. >> Mhm. >> deterministic, well, again, thinking about it in an MDP-ish kind of way, there's no stochastic transitions in this particular. Picture. >> Right. So if I'm in, state 2 and I go right, I always end up in state 4, period. >> Right. >> Mm-hm. >> Game. I guess, a game is because it's more than one player? >> Sure. >> Of perfect information. That doesn't quite sound like the same terminology that we used in the MDP setting. But I'm wondering if that's like, I know what state I'm in, when I'm making a decision. So, it's like an MDP as apposed to a POMDP. >> Well it, that's exactly right. It's, it's that you know what state you're in and. Yeah. That's exactly what it means. It's like being the MDP versus the Palm DP. That's a great analogy. >> Cool. And does it matter that it's a tree like this? because when we were looking at MDPs, we had more complex structures of graphs and things. >> Well, you can think of this as unrolling the MDP if you want to. >> So then those states are sort of time stamped and history stamped. For, yeah, for the purposes of this discussion, yes. And that's a perfectly reasonable way of thinking about it. >> But, okay. >> But in general, we're going to be thinking about game trees, but actually, we're going through all of this for nothing, because we're going to discover pretty

soon that none of this matters. >> [LAUGH] >> but, give me a couple of slides to get there, okay? >> [LAUGH] Sure. >> Okay. So this is about the simplest or at least the, the least complicated game that you can think about. Two players, zero sum. Finite deterministic game of perfect information. You know, basically, I can look at this tree, I know everything I need to know, and I can make decisions about what action I might want to take in order to maximize my reward. Okay? >> Good. >> All right. Now, in MDPs of course we had this notion of policies, right. You remember what a policy was Michael? >> Mapping from states to actions. >> So in the game theory world, we have something very similar to policies. We call them strategies. So, all a strategy is a mapping of all possible states to actions. So for example, here's a strategy that A might have. When in state 1, go left. And when in state 4, also go left. >> Seems like a terrible strategy. Does it? >> Well, just, if nothing else, just in state 4. >> Sure, but it's a strategy, right? >> Okay, but it's just, it's a strategy for one of the players. >> Right, exactly, each player has a strategy. And that makes sense, right? Before when we talked about a policy, and mapped [UNKNOWN] to action, there was only ever one player, only ever one agent. And so we didn't have to worry about what other strategies there were. Here, when we talk about a strategy, it's always with respect to one of the players of the game. Okay, so, question. I've just given you one strategy, which is what A does in all the states A could potentially end up in. How many other strategies are there for A? >> For A? Okay, that sounds like a quiz. >> That does sound like a quiz. Let's make it a quiz.

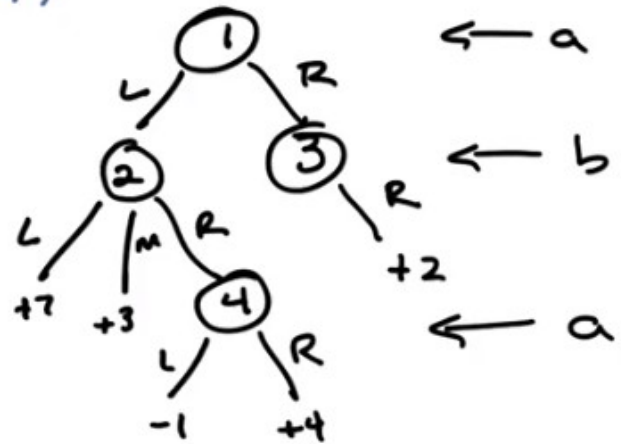
4. >> Okay Michael, so here's the quiz, because you made me make it a quiz, I decided to make it even harder, so I want you to tell me how many different strategies are there for A and how many different strategies there are for B. >> And you are talking about just deterministic mappings only, right? >> Well, we are in a two player zero sum, finite, deterministic game of perfect information. >> So, [LAUGH]. >> So yes, I mean deterministic. >> Okay, alright. because, you know, it could be that it might be helpful to be stochastic. >> Oh, that's true, that's true. So in fact, that's a good point. I was going to mention that later, but I'll mention it now. What I just wrote down here, is actually not just a strategy, it's something called a pure strategy. >> Hm, it's always about purity with you. >> It is, purity of chocolate and bacon mainly. But, yes. So, these are simple pure strategies. Okay, so how many pure strategies are there for A and B. >> Okay. >> Okay, then let's go.

A SIMPLE GAME (THEORY)

2-player zero-sum finite
deterministic game of
perfect information

STRATEGIES (PURE)

A: $(1 \rightarrow L, 4 \rightarrow L)$



QUIZ:

How many STRATEGIES FOR:

A 4 B 3

leaves = a's reward
+ b's -reward

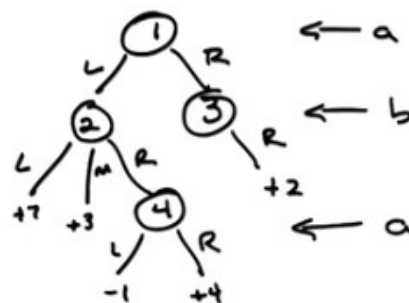
A 和 B 的俱體有哪些 strategies? 見下圖.

5. Okay, Michael, you ready? >> Yeah. >> Alright, what's the answer? >> So I was thinking about A before you threw in the B. So let me, let me not think about B yet. I'll just think about A. So you had said in state 1, it can go either left or right. And in state 4, it can go either left or right. So boy, that sounds a lot like two times two equals four. >> Yes that's exactly right. But generally speaking, so actually walk me through that again Michael. How did you get two times two? >> So, well I, I had a little choice about how to think about it. One is that in some sense, if you go right from 1, then you don't really have to make another choice. >> Right. >> But if you go left, then you have this other choice to make of either left or right. So it's, you if, if you're just writing it down as a mapping from state to action, you've got two choices at state one, and two choices at state four. >> Mmhmm. >> And so that is two times two, right? You can make, independently choose each of those. >> Right, that's exactly right. So in fact it's important there that even though if I can gone right on one, I would never, have to make another choice because I can't reach state four. In order to be a strategy, you have to basically say what you would do, in all states where you might end up. >> Okay, that's fair. >> Okay, alright so what about B, using that incredible reasoning? >> [LAUGH] So yeah, so B seems a little trickier because here it can only ever matter whether you're in like if player A sends us down to the left then we have a choice of three. If player one sends us down the right, we have a choice of one, which is really no choice at all. >> Mmhmm. >> You can have any color car you want as long as it's black. >> Yeah, like my Tesla. >> I was thinking the Model T, but maybe T stands for Tesla. >> T does stand for Tesla. >> So, by the definition of how many different you know, sort of reachable strategies it would be one answer but, if you're defining it the way you're defining it, it's going to be three times one or, three. >> Yeah. And that's exactly right. Good Michael. So, let's actually write down what those strategies are.

A SIMPLE GAME (THEORY)

		B:			
		②	L	M	R
		③	R	R	R
A:	①		7	3	-1
			7	3	4
			2	2	2
			2	2	2

④				
L	L			
L	R			
R	L			
R	R			



上圖 A 中的

① ④

L L ← 表示球在 ①處時, A 讓球往左走; 球在④處時, A 該球往左走.

B 中的

② L

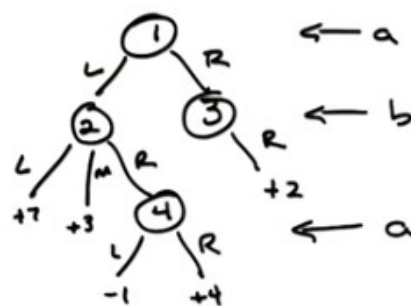
③ R

表示球在 ②處時, B 讓球往左走; 球在③處時, A 該球往右走.

6. Okay, Michael so I have written out based on upon your rather impressive way of thinking about it. All the possible strategies for a. And all the possible strategies for b. A, can go left, left, left right, right left, right, right. For states one and four respectively. And b can go Left right, middle right or right right. >> Right. >> Right. In particular, three in stage three we can only go R. Which makes me a? >> Conservative? >> A pirate. >> Oh. That makes more sense. >> It does make more sense. Okay. So, now why did I write it this way Michael? Well, I wrote it this way because if I write it this way with all the choices, all the strategies that b has up here and all the strategies that a have here. I actually form a matrix. And I can put in each of cells of the matrix the value of taking a particular strategy from A and a particular strategy from B. Does that make sense? >> Yeah, that's very clever. >> Yes, it is very clever. I'm very happy that I came up with it entirely on my own. Okay, so let's start filling in these numbers. Or, instead of filling in the numbers ourselves, we could ask the students to do it by making it a quiz. >> Nice, I see. >> Shall we do that? >> Sure doesn't seem very hard. >> Okay, so let's make certain everyone here understands and what exactly we're asking you to do. We're saying that if for example A takes this first strategy go left and stay one and left and stayed four. And B takes it's first strategy which is go left and stay two and right and stayed three. What is the value for A that will result? So, let's actually do the first one as an example and ask everyone to do the rest of them. That seem fair? >> Yeah, that's what I was going to suggest too. >> Okay, good. So let's see. If A chooses to go left in state one, since A goes first, we'll end up in state 2, right? And then in this first strategy. B goes left in state 2, which means we will end up going down this path and the value there is +7. So

seven is in fact the value of this game with respect to A. Now we know that because this is a two player zero sum finite deterministic game of perfect information. That if A gets a value of 7, B gets a value of -7. So we could write -7 here, seven comma minus seven here. But since we know that it's equal and opposite let's just write down the value for A. Okay? Just for compactness-sake. That seem fair to you? >> Yeah. >> Okay, cool. So with that in mind, let's see if we can fill out the rest of this table. Think you can do it? >> I think so. >> Okay. Then. >> Can you give me a place to write everything? >> It will magically appear. Go.

A SIMPLE GAME (THEORY)



		B:			
		②	L	M	R
		③	R	R	R
A:	①				
	L		7	3	-1
	L		7	3	4
	R		2	2	2
	R		2	2	2

(本段圖跟上段一樣)

A 選的時候, A 要將每行都試完, 然後選結果最好的.
B 選的時候也一樣.

7. Alright Michael. Tell me the answer. Lets go across. >> Alright. Across. So. The first row it's all left, left, but the second left doesn't matter because that's in state four. Oh it might matter. I take it back. So left, left you did was seven. Left and then, player two goes to medium, middle. >> Middle. >> That would be three, so then the pay off is three. >> Mm hm. >> If the first player does left, left and the second player does right, right; then we're going to go left, right, left again, or for a minus one. >> Mm hm. Oh wait, hang on. Left, right, left. Yes, huh, yeah. The next row should be exactly the same as this one, except for in that third case. So it should be seven, three. And the r, the difference there. The r in four only matters in this case where we've gone to the right state in [INAUDIBLE], sorry, right action from state two. And so that should give us a plus four, so four. >> Correct so far. >> All right half way there, so, all right now, now, we've got something different so now player one is going to the right, so now actually player one's choice in state four never matters, so the next two rows are going to be identical to each other. And so in the first case, oh, actually [LAUGH] all of these numbers are going to be identical to each other, so player one goes right, player. Or, sorry, player A goes right, player B has to go right. There's no choice. So we're going to get 2s, no matter what now, so it's going to be two, two, two, and then the second row will repeat that two, two, two. >> Okay. So there you go. Michael,

you are correct. You can feel very good about yourself. Very good, very good. So you're right, Michael, we do have we do have this nice little matrix and we have these numbers, and you know what's really interesting about this? That we can figure out what the payoffs are for B without having to do any additional work. >> That's true. But it's even more interesting than that. Which is, now that I've written down this matrix, nothing else matters. This whole tree, all these rules. None of it matters (然後把那個樹擦掉了). It's irrelevant. >> Wait, wait, but. Okay. >> Because everything about the game is captured here in this matrix. This is actually called the matrix form of the game, and it comprises everything you need to know about it. It doesn't matter what it means to go left or right or what it means to go middle or whatever. The point is by following this strategy and this strategy, you end up with 7 for A. This strategy and this strategy, you end up with two for a, period. And how you got there does not matter. >> But it does creep me out a little bit that what your saying is that all that matters is the matrix. >> You know we should write a movie about that. >> I think it's been done and it's scary. >> Well, not the first one. >> It was kind of scary. Like people are trapped in this big box and their slurpyness. >> But their happy. >> But they don't know their happy. >> What do you mean they don't know their happy? >> [LAUGH] >> What is means to be, Michael, Michael, Michael. >> Speaking of which now how do we, how do we figure out. What do we do with this now? >> So that's the big question right? So, what was the whole point of doing reinforcement learning? The whole point of doing reinforcement learning was to optimize your long term expected reward right? >> Yeah. >> And so you pick the policy that would get you to the best place possible, so here's a question I have for you. Give that this is everything we need to know, this little matrix of values. These are all the policies A could choose from. We're calling them strategies here. But, you know, strategies, policies. There are four A can choose from. And there are three B can choose from. What will you think A and B actually do? >> Okay, so, you know? A wants the rewards to be high. So, 7 is the highest, so A should choose the upper left corner. >> But A can't choose the upper left corner. >> Why? >> A can only choose a strategy. B gets to choose some strategy. >> I see. So A is choosing the row and then B gets to choose the column. >> Yeah. >> So then B should choose that also because that's what A wants. >> No, B wants to maximize what B gets. And remember, B always gets the opposite of A because. Because it's a two-player, zero-sum, finite game of perfect information, deterministic something. >> So, so you're saying that if we, if A chooses the first row. >> Say. Mm-hm. >> Which is the left, left strategy, and B now has a choice between three values and will choose the one that is worst for A, which would be the -1, so that would be a terrible thing to do. >> Yep. >> Okay then, so then what if A chooses the second row? >> Okay if A chooses the second row? >> So now again B is not going to let them have that seven which is kind of sad, but still cant make it to bad for A so B would choose the M R strategy. >> Mm-hm. >> And then get, and then that would be a 3 for A, a -3 for B. But wait, hang on. So that was done thinking of it as A move, A kind of making the choice first. >> Mm-hm. >> That doesn't seem fair. >> Okay, well then do it the other way. >> Alright, so if B chooses first, then B could choose the far right column because that's where the -1 is. >> Mm-hm. >> That's where it's going to be happiest. See, B seems kind of, kind of mean. It's like it's happiest when others are suffering. Well, but A is also happiest when others are suffering. >> I see. So, if B chooses that column then A wants to choose the second row and get the 4, so B could choose the middle column, which then A would choose the two. >> A would choose what? >> One of the twos. One of the bottom two rows. >> No he wouldn't. >> Oh, right. A is trying to maximize, so A would choose one of the top two rows. Oh, yeah, that makes more sense. >> Right. >> Then choose a three which is better for B than having chosen the far right, and B could choose the far left column, but then A would choose one of the 7, and B would be unhappy with that. So we'd end up with B choosing the middle column and A either choosing the top or the second row. Mhm. >> So that's kind of the same answer we got the other way. >> Yep. >> Huh. >> That's exactly right. >> And is that, was that just luck, or did you make this example to make that happen? >> No, I didn't make this example to make that happen. In fact, the process you went through is exactly the process you would expect to go through, and you will always end up with the same

answer. For a two player zero sum game. You know, deterministic, finite, all those other words. >> [LAUGH] >> The process you just went through, is exactly the process you would expect to go through. So let's see if we can, if we can be a little bit more, you know, explicit about the process you went through.

A SIMPLE GAME (THEORY)

MINI MAX

		B:				
		②	L	M	R	
		③	R	R	R	
A:	①					
	L					
	L					
	R					
	R					
	④					
	L					
	R					
	L					
	R					
			7	3	-1	
			7	3	4	
			2	2	2	
			2	2	2	

→ MAX
A considers
worst case
counter
B considers
worst case
counter
→ MIN

8. So, in particular, the way I heard you write it down was that A must consider the sort of worst case counter strategy by B right? >> I see, because when A chooses the row, then B was going to make things bad for A along that row, so that's the counter strategy you mean. >> Right, and in fact, when you try to do it the other way with B. Well, B has to do the same thing. B has to consider the worst case counter, as well. And in this particular case, the way we set it up. Where the values for A. A is always therefore trying to maximize. And B is always trying to minimize A. Which works out to be the same thing as maximizing itself. Does that make sense? >> Yeah! I mean, other than the fact that, you know, I name my first child Max. I really wanted to name my second child Min. >> [LAUGH] That actually would have been pretty cool. Why didn't you do that? >> Because I'm married to someone with more sense than I have. >> Yeah I understand, I completely understand. Okay so, A is trying to maximize, B is trying to minimize, they both have to consider the worst case that the other will do. And so what's going to force them through exactly the same process you went through. We just figure, I'm going to make the choice so that my opponent makes the counter choice, the worst case choice, I will end up as best as I can. So A is going to basically try to find the maximum minimum, and B is trying to find the minimum maximum. >> Hmm >> In fact that strategy has a name, or that way of thinking about it has a name. What do you think it's called? >> The minimum, maximum. >> Yes, or Mini max. >> Which is movie production company, I think. >> No, that was Miramax. Do you recall that where you have seen Mini max before Michael? In some other class that you once taught or once took years and years ago? >> Mmm. No. >> Mini max was exactly the algorithm that we use for game search. And intro to AI. >> Oh, which was a game tree, which is just what we started with in this case. Even though we turned it into a matrix. >> Exactly. So in the end, we, you know, this matrix induces a game tree, if you want to think about it that way. Or, game tree induces a matrix, if you want to think about it that way. And the strategy then, in sort of basic AI search, and the strategy in game

theory is Minimax when you're in a two player zero sum game of perfect information. >> So is there a way to do alpha beta pruning on this? >> All alpha beta pruning does is gives you a more efficient way of finding the answer. >> I see but it's the same answer no matter how you set it up. >> Right. Cool. >> That's right. Okay, so this is pretty cool. So, we have set up a kind of game where we have multiple agents, you know, or at least two agents in this case, who have different strategies. And we actually sort of know, [SOUND], you know, sort of, in a world where it's a zero sum game, and you know the other person is trying to minimize what you get, maximize what they get. That the mini-max strategy would actually give you sort of an answer, in this case, by the way. We say that the value of this game for A is three. If A does the rational thing, and B does the rational thing, that is trying to maximize their own value, you will end up in this situation. That's kind of cool, don't you think? >> Very cool. I feel like there should be a theorem. >> There is in fact a theorem. I'm going to write it down.

FUNDAMENTAL RESULT

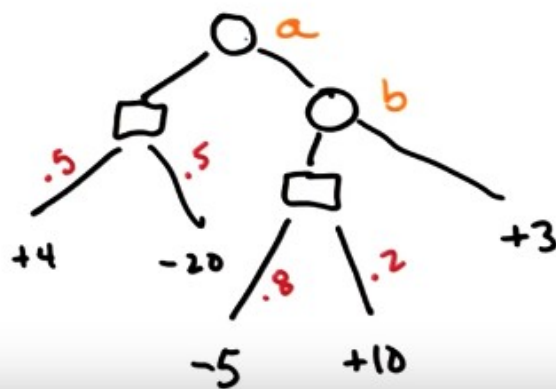
In a 2-player, zero-sum
deterministic game of
perfect information

Minimax \equiv Maximin... and there
always exists an optimal pure
strategy for each player

9. Okay Michael, so here's this theorem that you, you so desperately wanted. You ready? >> Yep. >> I'm going to read it to you, because you can't read. In a two player zero-sum deterministic game of perfect information, minimax equals maximin. >> Alright you told us what minimax was, but you didn't tell us what maximin was. >> Well maximin is like minimax, except the other way around. So a is trying to. [LAUGH] >> You know, one side is trying to minimize the maximum, the other side is trying to maximize the minimum. >> Okay. >> It's exactly what we described before, just depends upon whether you're looking at it from a's point of view or b's point of view. >> Oh, I see, like, which, do you choose a column first or do you choose a row first? >> Exactly, so whether you go A first followed by B, or B first followed by A. You're going to end up in the same, with the same result. And that, more importantly, or at least as important, there always exists an optimal pure strategy for each player. In other words, you can solve those games and you know what the answer is. Once you write down the matrix. You just do Minimax, or you do Maximin and you end up with the proper answer. And now you know what the optimal players would do. Now there is a subtlety here which I got it a little bit in the previous slide, when I talked about rational agents. And what we're sort of assuming in everything that we discuss here is that people are always trying to maximize their rewards, okay? So we define the, the reinforcement learning problem that way. That my goal is to find a policy that maximizes my long term expected reward. You know so I'm trying to find, to get the best reward that I

can. And what you're assuming here is that everyone else is doing the same thing and they're assuming that everyone else is doing the same thing. >> Okay. >> Does that make sense? >> It does thought I'm a little bit stuck on this word optimal at the moment. Right. Well, that's what I'm trying to get at actually. That optimal here really means I'm maximizing the reward that I can get, and I'm assuming everyone else is doing the same thing. And I'm, furthermore, I'm assuming that they're assuming that everyone else is doing the same thing. >> So, so I guess I'm wondering whether. Whether this theorem is vacuous in a sense that are we defining optimal to be mini max. >> What we're defining optimal to be, I think. So that's a good question. I think I would unroll the vacuous one level by saying this. Optimal here, basically has to be optimal in respect to what? And the respect of what here is the underlying assumption that everyone is trying to maximize their rewards. And that everyone knows this. So, [in a world where you have perfect information. It's zero-sum. Then the strategy of Minimax and Maximin give you the same answer.](#) And that furthermore there is always some place where the column and the row cross, the best column and the best row cross. And that is always going to be the solution to that particular game. Now, if we weren't in a two-player zero-sum deterministic game of perfect information, that might not be the case. But in a case where we're in this sort of simplest version, where everyone's being. Rational, that is, optimal, that is, trying to maximize their, their own reward. And assuming everyone else is maximizing their own reward, this is the right thing to do. Now, I've got a little weasel word here as well which we're going to get to in a moment which is this notion not just of a strategy but of a pure strategy. [INAUDIBLE] There's a reason why we have notions of pure strategies because in the end as we get more complicated we're going to have to do it with impure strategies. >> Mmm. >> Okay, but are you with me so far? >> I think so, yeah. >> So basically all that stuff we did in AI with game trees and search is kind of what you would expect people to do if they knew everything. [LAUGH] >> So, So, I feel like I could prove this theorem in the case of trees because you just propagate, you just kind of commute values from leaves up to the root. >> Yeah. >> And, it, it doesn't matter. There is no notion of who goes first or who goes second. So there's just going to be one answer, to that process. It's not obvious to me, how to show it, if you, once you've. Turn the tree into a matrix, that that matrix, I guess because it captures the same information, it ought to be the case that this is still true, but like, I'd have to kind of sit down and think it through. >> No, and it's, so, so, to help you think it through, I guess what I would suggest is if you have the matrix, you can create a tree that's consistent with it. Because every row and every column represents a strategy. You don't know what that strategy is, but you can, since it's a finite matrix. You can construct a tree that is consistent with that matrix. In fact, possibly an infinite number of them, I'm not sure, but you can certainly construct at least one that is consistent with it. And then once you have the tree you just do what you said. >> Good. >> Alright, good. So we've got this fundamental result and now what we're going to do is we're going to try to be a bit more interesting. But it is important to go through this because now we've got some basic vocabulary and some basic building blocks okay? >> Yep. >> Alright.

2-player, zero-sum
~~no~~ deterministic game of
 perfect information



10. So here's another game tree. We have, again, two players, A and B, and A gets to make a choice, first, to go left or right. And b will find herself in a situation, perhaps, where she gets to choose to go left or right as well. I've drawn these little square boxes, though, to represent chance. So what this means is that if A goes left, you end up in a chance box where you flip a coin and 50% of the time you end up over here (+4 那裡), and 50% of the time you end up over here (-20 那裡). Along a similar vein, if a goes right and then b goes left, then you end up in another chance node, and 80% of the time you end up here (-5 那裡), and 20% of the time you end up here(+10 那裡). Alternatively, if b goes right, you always end up here. Does that make sense? >> I think so. Uh-huh. >> So what we've done is we've gone from a two player zero-sum deterministic game of perfect information to a two player zero-sum **non-deterministic** game of perfect information. Okay, so we've relaxed, we gone at least one level up in complexity. Okay, so do you understand this tree? Do you understand this setup? >> Yeah, I think so. I mean here the stochasticness is happening essentially at the leaves. >> What does that mean? >> That there's no choices to make for either player after this randomness happens. But is that. That's not what you mean in general, right? >> No, that's right. There could be, after here, you end up in a different state where you can then make more choices. But because I don't have enough room, maybe because I want to make it simple, it just sort of ends after that. >> Okay. >> But yes, this tree could keep going on and there could be choice, random, choices chance nodes happening everywhere. There could have been a chance node that happened at the very beginning, even. >> I feel like I could work out the value of this game. >> Oh, well, how would you go about working out the value of the game? >> I would probably have it as a quiz. >> Okay, but what would the quiz look like? >> It would say, what's the value of this game? >> Okay. Do you think that just having that be the quiz is the best way for someone to learn or do you think maybe it can be done in stages? >> Oh, I don't know. Do you have other questions that you want to ask? >> Well, you know, how would you go about determining the value of the game? I think the first thing that you would do, at least if you were patterning after what we just did, is you would try to write down a matrix. >> Sure. >> So, I'm going to write down a matrix. So that our students will get a chance to, to work out what that matrix looks like and then from there figure out the value of the game. Are you okay with that? >> Sure. >> Okay. So if we were going to do that of course. The first thing we'd have to do is we'd have to figure out, figure out what the strategies are. So. What are a's strategies. >> I would have called them left and right but it, but they're not labeled. >> Yeah well let's do that we can call them left and right. A can go left or A can go right.

What about B's strategies? >> B can go left or right. >> And B can go left or right.

11. >> Here's your first quiz question. In a world where A can go left or right and B can go left or right. What are the values that you would put in the cells of this matrix? Again, this is zero sum. So, these values are from A's point of view and implicitly, there's we know what the values are for B. So, we're just looking for the values from A. Okay? Do you understand the quiz, Michael? >> Yep. >> Okay, so go.

2-player, zero-sum
nondeterministic game of
perfect information

		B	
		L	R
A:	L	-8	-8
	R	-2	3

12. All right Michael, you know the answer? >> Yep. >> Okay, what's the answer? >> Do you want to know the value of the whole game? >> No. I want to know the matrix. >> [LAUGH] Okay. So right. So, if A goes left, it doesn't matter what B does. And at that point there's a chance node, and it's 50%, 50%, 4, -20, which I feel like ought to be, negative eight. >> That's right. How'd you get negative eight? because half of 2 is sorry, half of 4 is 2 and half of -20 is -10. So -8. >> Right, so you just took the expectation of where you would end up. >> Exactly. >> Okay. What next? And that, it doesn't matter what b does. So then negative eight is also in the upper right corner of the matrix. >> Fair enough. >> The next easy one to do is if both go right. A goes right and b goes right then we get the three. It's just sitting there. >> Mm-hm. >> And the last thing requires me to do some multiplication. So >> Rrr. >> -5 times 0.8 is like -4. >> Mm-hm. >> 10 times 0.2 is like 2 so -2. >> That is correct Michael. >> Shoo. >> So now we have the matrix. Now, what did we just notice here? Remember what I said about matrices before? >> It has all the information that you need. >> Right, so in fact none of this matters (然後把那個樹擦掉了). Who cares how we got here? >> wait, but, but, oh. >> [LAUGH] I love erasing stuff! >> So here's the thing, we cannot reconstruct that tree from this matrix. >> We can't reconstruct, well, actually we could reconstruct that tree from the matrix. >> No, we can't. >> Yeah, we, of course we could. >> No, we can't. >> Yes, we can, because there's an infinite number of trees we could do, and one of them is that one. >> I see. >> We just don't happen to know that it's the right one. >> I don't think that's really what people mean when they say I could reconstruct this. Like I could reconstruct the crime at this crime scene, it could be any of a million things. Like no, we can't construct that specific tree. >> But you know what it doesn't matter, because the only thing that matters is the matrix. >> Ahh. >> And you will notice, Michael, that you did all that multiplication and you multiplied 0.8 times 5 and after a couple of seconds of thinking about which was probably edited out but I know how long it took you to do. You came up with the right answer. That's great. But notice that the number you came up with doesn't say anything about expected values and what the probabilities are. It doesn't even matter that it's nondeterministic, because once you have these numbers, these expected values, that's what you're going to end up with. So who cares what the original tree was? Who cares if you can reconstruct it? Who cares what the rules of the games are? All you know is I have a choice of two strategies. We call them left and right here, but we could have called them one and two,

or Q and Z. It doesn't matter. For the purpose of solving the game. [So what is the solution for the game by the way?](#) >> Oh, A is trying to maximize right? >> Yes. >> So A would never, ever, ever, ever want to go left. >> Mm-hm. >> So A's going to go right and then B is trying to minimize. So B is going to go left and [we get the -2.](#) >> I believe that is correct. And it makes sense the other way as well, right? That B would not ever choose to go right because then A would choose to go right as well. >> Gotcha. >> So you'll still end up here.

2-player, zero-sum
 non-deterministic game of
 perfect information

THAT OTHER THEOREM
 STILL HOLDS!

VON NEUMANN!

		B	
		L	R
A:	L	-8	-8
	R	-2	3

13. >> So, [here's in fact another theorem](#) for you Michael. So, [it turns out the theorem that we wrote down before is still true in the case of non-deterministic games \(another theorem 應該就是指的這個\).](#) > > Oh, cool. >> Of perfect information, right? By the way, do you know whose theorem this is? >> Charles' theorem? >> No, although that would be cool. >> Von Neumann's theorem? >> Yes, [this is von Neumann's theorem.](#) >> Oh really? >> Yeah. Did you just guess? >> Von Neumann, he's responsible for everything. >> Do you know, well, you going to remind everyone who von Neumann is? >> He's my uncle? >> No. >> No, you're right. You're right. Von Neumann, so we talk about von Neumann architectures in computer science, that the basic design of a microprocessor is still following the ideas that that he worked out. >> Right, the von Neumann Machine. So there you go. So, this is pretty good, right? So, what we did, so what did we learn so far? [What we learned so far is, the only thing that matters is the matrix.](#) And once you have the matrix, you used mini max, at least if you're in a two player zero sum game of perfect information. At least if you're in a world of two player zero sum games of perfect information. We can just write down this matrix, throw everything else away and use mini max or maxi min and figure out what the value of the game is, which is the same thing as saying as, we know what the policy is. The kind of joint policy, in a world where everyone is being rational and trying to maximize their rewards and assumes everyone else is doing the same. That's pretty cool, I think. >> Awesome, so though, you know, I feel like I could make a matrix that wouldn't have this property. This maxi min equals mini max. And, and that we couldn't make a tree out of it. >> You probably could, but I'm going to guess that it's going to require that it's no longer zero sum. >> Nooo, I'm going to say no to that. >> Well, you know what, I, actually now that, I think I understand what you mean and the answer is yes you could. And in fact, that's what we're going to do next when we

relax the next little bit. >> Ooh. >> So, would you like to do that? >> Sure, though I'm afraid if we get too relaxed, we might just fall asleep. >> In fact, why don't we do that, why don't we take a nap and then come back? >> [LAUGH] the, the joy of being on video, asynchronous napping. >> Done. Okay, I'll see you in a second.

2-player, zero-sum
non-deterministic game of
perfect information
hidden

Mini-Poker

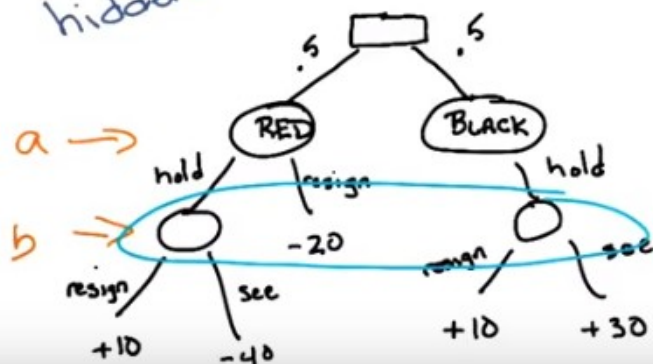
- A is dealt a card, red or black 50%
- A may resign if red: -20 cents for A
else A holds
- B resigns: +10 cents
- B sees:
 - if red: -40 cents
 - if black: +30

上圖中 A 是一開始就看了 card 的顏色的, B 一直不知道 card 的顏色, 直到出最後結果時 B 才有可能知道. 下段的圖更清楚.

14. Okay Michael, so here is a, another little game for us to play. And, what I want you to notice about this game before I describe it to you in detail is that we have relaxed yet another one of the constraints. So, we started out playing two player, zero sum, deterministic games of perfect information. There's also a finite in there somewhere. From now on we're just going to assume everything's finite, because why not? And then, [what we did last time](#). Just a few seconds ago. [Is we relax the deterministic part, so we have two player, zero sum, non-deterministic games of perfect information, and now we are going to relax the requirement for perfect information.](#) So now we're going to look at two player, zero-sum, possibly non-deterministic games of hidden information. And this is really important Michael because, [this last little bit of relaxation, going from perfect information to hidden information is going to be sort of a quantum leap into the difficult problems of game theory.](#) So this is where it actually starts to get interesting [so we have been building a foundation so far and now we are going to get to the interesting and complicated stuff](#), okay? >> Wow, one of the things that I learned just now is that the opposite of perfect is hidden. >> Yes. >> I always thought the opposite of perfect is imperfect but okay but hidden. I guess if I do not have a perfect face I should hide my face. >> That's in fact the atomology of the phrase. >> Alright, I understand now. >> Yeah, cool, It's in wikipedia. Here we go, [let me describe the game to you](#), are you ready? [This is a version of mini poker where there are a set of cards, but they have no numbers or faces on them, they are just red or black, okay.](#) >> [And red is bad, for our hero, Player A. And black is good for our hero, Player A.](#) >> Okay? >> I see. >> So-- >> Wait. Why's it have to be red? >> I don't know, man. You know, red. You know how it is. Okay. So, here are the rules. They're all written down on the screen, but let me walk through them with you. So [A is dealt card](#)

(分牌給某人). Magically. It will be red or black. And the probability of it being red or black is 50% each. Okay? >> Yes. >> Right. So we have a uniform prior over, over the color. Now, remember red is bad for A and black is good for A. So, it's going to turn out without loss of generality, that if A gets a black card, A's definitely going to hold onto the card. Okay? Now A gets this card. B, player B, does not get to see the card. A can choose to either resign(放棄, 即看了牌的顏色, 而把牌扔掉, 誰也不給, 下段已確認) or to hold (即看了牌的顏色, 而把牌 offer 給 B, 但不告訴 B 顏色, 問 B 要不要, 下段已確認). If A resigns given a red card, then he loses 20 cents. Okay? >> Okay. Wait. So A's dealt red. A may resign if but only if red. >> Right. >> And then A loses 20 cents. >> Right. >> Okay. Alright, okay. >> Okay, so this is a betting game. It's not strange it makes perfect sense its sort of a metaphor for life. Now A can choose to hold instead, hold the card. Thus requiring B to do something. So if A holds the card B can either resign or can demand to see the card. Now if B resigns then A gets 10 cents. Regardless of the color of the card. Okay? >> Yep. >> That make sense? >> Yep. >> Okay. Now if B chooses to see the card, in fact demands to see the card. Then if the card is red, then A loses 40 cents. But if the card is black, then A gets 30 cents. And since we're betting, this means that whatever A wins, B loses and vice versa. That makes it zero sum. >> Got it. >> Okay, is this all make sense? >> Yeah, I don't know if I can hold all those different numbers in my head. But I, but the basic pattern of it is, that as you say Red is, red is bad, black is good. If A gets a bad card, A can essentially either fold >> Mm-hm. >> Right, resign? Or kind of bluff, like hey I've got this great card and then A and if B believes that the card is bad then and calls A or, or, or, and, and just folds then, then A gets, A wins that. But if B says no I think maybe you're bluffing and calls him, then everybody's rewards are more extreme, I guess. >> Exactly. So it's just a version of poker. A weird version of poker. Simple version of poker. But a version of poker, none the less. There's a minor little detail here which isn't that important, but you know, notice it's written that A will A may resign if its red. Basically, A will never resign on a black card. Because it just doesn't make any sense. And so there's really, it just, there's not point in riding it out. Okay. Because black is always good, sort of, nothing bad can ever happen to A, if A gets a black card. So there's really sort of no point in riding anything out. But that's just a minor detail. Regardless these are the rules. Okay? >> Okay. >> You got it? Sure. >> I'm going to re-draw this as a game tree which might make it a little easier to keep all the rules in your head.

2-player, zero-sum
~~non~~ deterministic game of
~~perfect~~ information
 hidden



MINI-POKER

B:

resigner seer

A:
 resigner
 holder

15. Okay so Michael [here's the tree version](#), of what I just said. So remember I draw squares as chance nodes. And so, chance takes a, chance. And half the time we end up in a state, where I have, where A has a red card. And half the time I end up in a state where A has, or we end up in a state, where A has a black card. Okay? If A has the black card. Then B gets to choose whether to hold or not, so let me add a little bit of color for you since you wanted some color. This is the place where A gets to make a decision. Okay. Yep. And then this is the place where B gets to make a decision. Got it. Okay? So, A is going to either be in a red state or a black state. Only A knows this. B does not know this, does not know what state He is in. And so let's say A is in a black state. Well, A can only hold in that case. In part because it makes no sense to resign. And then B gets to decide whether to resign, and therefore A gets ten cents or to see, in which case A gets thirty cents. By contrast, when we're in the red state A can either hold or resign. If A resigns he loses 20 cents. If A holds then B gets to decide whether to resign and which case A gets 10 cents, or to see the card in which case [A loses 40 cents and B of course gains 40 cents](#). So this is just a tree version of what I just wrote. Does that make sense? >> Okay, yeah I can see how this kind of captures the flow of information but I feel like there might be a missing constraint to it. So [INAUDIBLE] I don't know if there's actually a constraint in this thing but [let me just point out something. And that is that B has no idea which of these two states \(藍圈中的那兩個圈\) he's in, and its hidden information](#). >> And because B doesn't know what states he's in. He doesn't know whether resign or see. In particular B knew that he was in this left mode state then he would always say see. If B knew he was always in the rightmost state, then he would always resign. But he doesn't know which one, so it's not entirely clear what to do. >> Neat! Where did you get this game? This game is awesome! . >> I got this game like I got all of the examples that we're using today, from Andrew Moore. >> He's clever. >> He is clever. And I have shamelessly stolen... All of his examples and materials for this. But he said it was okay, by putting up slides in the world and saying, feel free to steal all of the stuff. >> You want to write his name so that people can credit him? >> Yeah, I'm going to write it at the very end. When we talk about. >> Okay. >> What we've learned. >> Awesome. >> Okay. So here's a question for you Michael. We know that I wrote down a bunch of words which describe the game, and then I made it a tree, because I could do that. And it makes it nice and easy to see what's going on. But now we know, that at least everything else we've done. We want to make a little matrix. And so we want to figure out what the value is for different strategies for A and B. So I'm going to assert, and I think it's pretty easy to see, I hope. That A basically has only two strategies. Either A is the type of person that resigns when a card is read. Or A is the type of person who holds, when a card is read. Agreed? >> Interesting, okay. Yeah. Right. So it really is a conditional policy, right? It's basically If red hold to resign, if black hold to resign, but your point is that black isn't really a choice, red is a choice, and there's only two choices. Okay I can see that as the two strategies sure. >> Right and of course this does say, this is, this does kind of say when you're in the black case you know you're going to hold. So you know what's going to happen. Ultimately B can either be the kind of person who resigns, whenever A holds or chooses to see, whenever A holds. Right? >> I see so, like in the previous trees. B would have four different strategies Resigner see in the left state, resigner see in the right state. Here there's this kind of extra connection, this sort of a quantum entanglement between these two states that it, that makes them have to be the same. So there really is just those two choices. >> That's exactly right, although, I probably wouldn't have used the phrase quantum entanglement. But yes, there's an entanglement, certainly. because we can't tell which state we're in. You're either going to resign or you're going to see. And you just don't know what else to do. Okay so A's a resigner or holder. B's a resigner or a seer. [So the question is what numbers go in this matrix? And we're going to figure that out by having a quiz](#). >> I kind of saw that coming. >> Mm-hm, do you think you can figure this out? yeah, yeah. Sure. Sure. Yeah, I think so. >> Yeah, see, see, see. Yeah, I can do that. Okay, cool. So, let's go then. Go.

2-player, zero-sum
 deterministic game of
 perfect information
 hidden

pure strategy
 mixed

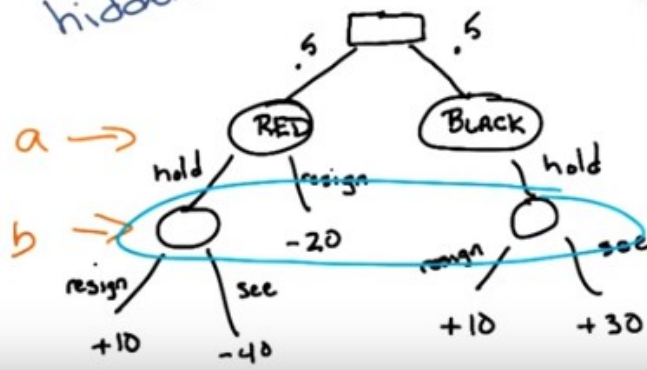
QUIZ

MINI-POKER

B:

resigner seer

A: resigner	-5	+5
holder	+10	-5



Minimax

Maximin

Full screen

16. Okay Michael what's the answer? Let's start with resigner resigner. >> Alright, resigner resigner. Resign resigner diner. Alright, E, so resigner vs resigner. So resigner, if A is a resigner that means whenever A gets a red card A resigns. Yes. >> And that would be a -20. >> Yep. >> But that's not going to be the answer is it. Because A doesn't always, A doesn't always get a red card. A sometimes gets a black card. If A gets a black card, then a resigner resigner that means B is going to resign and, and a +10 will happen (這其實是一條新規則, 前面沒講). >> Mm-hm. >> So those are the two possibilities and they're equally likely, so it's a -10 divided by 2 which is -5. Right. You were correct, sir. >> Whew. Alright. >> Which one do you want to do next? >> Resigner seer. >> Okay. >> So again, oh, so the, yeah, good. This is a good choice, because now it's the same as argument as before, except for when we end up in that far right node, and that means minus 20 in half the cases, plus 30 in half the cases Which is a plus 10 divided by 2, or plus 5. >> That's exactly right. Well done Michael. >> Thanks. >> Okay which one next? >> So holder resigner. So holder resigner. That means when A, gets a card. A is going to hold the card. >> Mm-hm. >> And that's true, red or black. >> Yep. >> And then, then, it's going to be B's turn, and B is going to, oh, we're doing holder resigner. So, it's going to resign. >> Mm-hm. >> So, oh, well, interestingly, I think that takes us to those two leaves, both of which are plus ten. >> Yep. >> Because, why does that make sense? Because B, oh, 'cause B doesn't get any, no. Right, right, right, because it's independent of the card. You actually said that when you explained the rules. >> Mm-hm. >> So its, its average of plus 10 and plus 10, which ought to be plus 10. >> It is in fact plus 10. Well done. Okay what about holder's here? >> Whew, alright, so that's a case where A holds so we go down those branches and we end up, we always end up in one of the blue circled states. >> Yep. And B sees half the time that leads to minus 40, half the time that leads to plus 30. So that's minus 10 divided by 2, which is minus 5 again? >> Yeah, that's exactly right. So that's correct, Michael. And that's pretty cool, isn't it? >> Yeah, I really like this game. I didn't think you could have anything that had sort of poker essence and be this tiny. >> So yeah, so we live in this really nice little structure here. So I have a question for you. You ready for the question? >> I'm trying to guess what it is, but sure. >> Okay, here's my question. What is the value of this game? >> I was thinking that you might ask that. So, can I, can I step through it, is that okay? >> Yeah, sure, go ahead. So a's choosing the first row or the second row. So if a chooses the first row, and then b chooses the

column, then if it's the first row, then b is going to choose the first column. So a's going to get minus 5. >> Mm-hm. >> The same story's going to go through on the bottom row. If a chooses the bottom row, then b's going to choose the seer position which gets the minus 5. >> Right. So from this so, it seems that the value of the games minus 5. But now let's do the same thing on the b column. So if B resigns, now A gets to choose resign or holder. And it gets a plus 10. And if B is a seer. Then A chooses between re signer and holder and gets plus 5. >> Yes. >> So then from this perspective, the value of the game is plus 5. So, so here's a case where it better not be that we could take a perfect information game and put it into a matrix and get this out, because this is something that can't, like, it doesn't, it doesn't fit your theorem, right? We can't get the value of it by doing minimax or maximin. >> Exactly, the problem here is that once you move the hidden information. Minimax is not necessarily, and in this case, definitely is not equal to maximin. So von Neumann. >> Fails. >> As we all see. >> Idiot. >> Yeah, what has, what has he ever done for us, anyway? His theorems and his computer architecture that rules the world. Anyway, so we seem to have a problem here. And the problem is that once we go to this hidden information, as I promise complexity enters in Michael, and now we can't do something very simple with the matrix, find a pure strategy that's going to work. It really is the case that A strategy depends upon what B will do and B strategy depends upon what A will do. And if you don't already know what that's going to be, you don't actually have a value for the game. And in some sense you can get every single one of these values. So, there's got to be some [INAUDIBLE]. >> I feel like, that's sort of what you'd expect in a game like this. Right. So, if, because of the way that it is. If you know that I'm always a resigner. That I'm always going to, what? [LAUGH] Oh, that when ever I have a red card, I'm going to resign, then. You know that if I don't resign, I have a black card, so you know that you should resign. >> Mm-hm. >> Yeah, so it's, it's, it's one of these things where if I am really consistent and never bluff, say, then you can take advantage of me, and vice versa. Like, if you always respond the same way to when I act a certain way, then I can manipulate that. So it's kind of like this game of this sort of mind game like I want you to not know that I know the thing that you don't know that I don't know. >> Right but the problem is that I know what you know and I know that you know what I know. And that I know that you know that I know that you know that I know that you know that you know what I know. >> Did you really think that I didn't know that? >> [LAUGH] And so you end up in this terrible situation. But. There was a key word that you used there, Michael, and it was consistent. And everything we've talked about so far, pure strategies, is exactly the same thing as talking about consistency. So, the way we're going to get around this is, we're going to stop being consistent, or at least, consistent in the same way. Okay? >> Yeah. >> And to cheat here, is that we are now going to introduce, instead of pure strategies. Let's see the opposite of pure. >> Is? >> Impure. >> Mixed, that's exactly right. >> Contaminated. >> No, so, rather than sticking with pure strategies, we're going to start using mixed strategies.

Quiz

B: resigner
A's expected profit:

$15p - 5$

B: seer
A's expected profit:

$-10p + 5$

$-5p + (1-p)5$

Mini-Poker

B:

	resigner	seer
A: resigner	-5	+5
holder	+10	-5

mixed strategy \Rightarrow
distribution over
strategies!

$P \leftarrow$ probability of being holder

Full screen

2:52 / 3:21

上面的框是因為 $10p - 5(1-p) = 15p - 5$

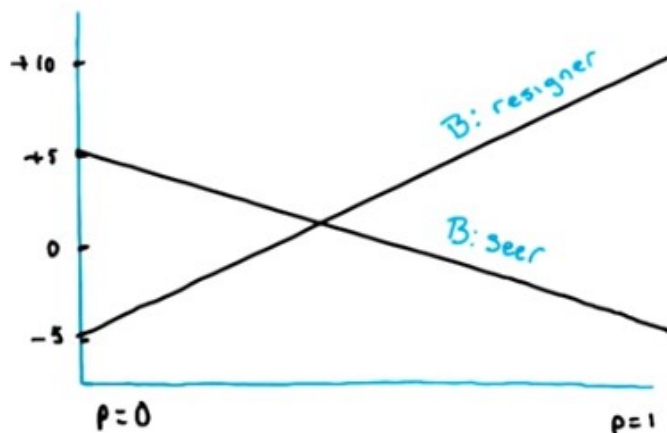
下面的框是因為 $-5p + 5(1-p) = -10p + 5$

17. So what's the difference between a pure strategy and a mixed strategy, Michael? Well, it's, it's simply this. A mixed strategy, simply implies, or means some distribution over strategies. So in the case of two strategies, like we have here, where you can either, A can be either a resigner or a holder, we're going to simply say that the mixed strategy for A is some value for P, which is the probability of choosing to be a holder. So do you, you see what's going on here? So the only difference between a mixed strategy and a pure strategy, is that for a mixed strategy, you choose some probability over all the different strategies that you might choose. So you decide that, you know, going into this I'm going to flip a coin. And half the time I'm going to be a resigner, and half the time I'm going to be a holder. Say. Or 30% of the time I'll be a resigner. And 70% of the time I'll be a holder. Okay? >> Yep. >> Whereas with pure strategies, you always chose on or the other. So technically, it's the case that a pure strategy's also a mixed strategy where all the probability mass is on a single strategy. >> Makes sense. >> So in this case we're going to, in fact, choose P to represent the probability for A of choosing to be a holder rather than a resigner. And so P can be 0%, probability zero, or can be probability one, or any value in between. You with me on that? >> Yeah, that's neat. >> Okay, good. To make certain you understand this I'm going to give you a little quiz. Which I have up here on the screen. You ready? >> Oh, I see it. [LAUGH] It's like those, those very square boxes. >> Yes. >> I didn't even realize what, what, what this could be about. >> It certainly couldn't be a quiz because Charles has never drawn a straight box in his life. I had drew those Michael. It took me 17 hours. >> Oh man, you are, you're committed to this and I appreciate that. >> I am committed to this. Okay, so, given that we have a mixed strategy, and we have a probability P of A being a holder, here's my question for you. In a world where B is a resigner, okay? B is always going to choose to resign. What is A's expected profit? To make it easy for you, I copied the matrix over here in the upper right hand corner. >> Wait, wait, wait. If B is always a resigner. >> B is always a resigner. >> Then, what's, and what is A? >> A is going to choose to be a holder with probability P. >> Oh, so you want this to be a function of P. >> Maybe. >> If it's, if it's, okay, sure, maybe [LAUGH]. It could, well, I mean, yes. It's a function of P, it just might

be a constant function that ignores P. >> It could be, in principle. Okay, now, [after you figure that out, I want you to decide, in a world where B is the seer](#), B always chooses to see the card. [What would A's expected profit be, in a world where A will choose to be a holder with probability P](#). Okay? >> Hm. >> Got it? >> And that's going to be, oh yeah, that could be different because it's a it's a different strategy, though you know seering. >> Yes >> Like like if you're a resigner you're a resigner, but if you're a seer then what you are doing is seering >> Mm-hm. >> That would be an anagram of resigner. >> How do you see these things? >> I don't know. They're just there. I just, they words just kind of mix themselves up. Alright anyway, I'm, I think I am ready to do the function. I'm, I'm ready to stop looking at the letters. >> Okay? And go!

18. What's the answer? If B is the resigner, we don't really care about the other column anymore. >> Mhm. >> Then what's going to happen is A is mixing between resigning and holding. >> Yap. >> And probability P is a probability of being a holder, so whatever P, whenever that event happens it gets ten, and whenever one the opposite of it happens it gets minus five, so I want to say ten p plus one minus p five. >> Okay, is that your. >> Let me try that again. >> Is that your final answer? >> Ten p. Well I could simplify it. >> Okay. Well then say it to me again. I'll write it out here so it's easy for you to see. Okay and that is. >> My answer? >> That's fair. Do you want to simplify it? You don't have to. >> Sure, let me simplify it, so it's a minus minus p five and a ten p so that's fifteen P minus one. >> Minus what? >> One. >> Minus what? >> Five. >> Uhun, there you go. Thank you. >> That's correct. We would obviously accept either answer or any combination of those letters. That, [LAUGH] >> No, I think I might have to do this quiz over actually. >> No, not the one. I'm talking about either 15 p minus five or ten p minus one minus p times five. Or ten p plus one minus p times minus five. Or any other combination that we can get push car to actually Bothered to, you know. >> Check. >> Im, Implement or check. Okay. Well, that was pretty good. And this, of course, is exactly the expected profit. As you put it, P times A as a holder and P times or P percentage 1 minus P percentage A chooses to be a resigner. And so it's just the weighted average between those two values. So, Kent, let me just double check that. So, if P is 0, that means it never holds, it means it always resigns and it gets -5, so that's right and if P is 1, means it always holds, so it should get a +10. And 15 -5 is 10. So, boom! >> Yeah, it works. You used math there, very good. Okay, what about B? >> B. So the same story, except on the seer side. >> Mm-hm. >> So yeah, I might need that space again. So 5, oh I see, right, minus 5 times p. >> Mm-hm. >> Plus 5 times 1 minus p. >> Mm-hm. If we simplify that we get. >> There's a minus 5 and another minus 5. So we get minus 10 p. >> Mm-hm. Plus 1. >> Plus what? >> Plus 5. >> There you go. See you learned. Okay you want to check it? >> yeah. Oh, that's a good idea. So again if P is 1, then that means you're always which should be the minus 5 and if we put in a 1 there, we get 5 minus 10 is minus 5. And if P is 0, that means that we're always a resigner, and we should get a 5 for that. And yeah, so we zero out the negative 10 and we get the 5. >> Exactly. >> Now, it's not clear to me why we're playing this game. Oh, it is clear to me why we're playing this game, because we want to figure out something about. Strategy that is mixed. >> Right. >> So this is how well a mixed strategy does, but not against another mixed strategy. This is against two deterministic strategies. >> But is it, Michael? But is it? So >> I'm going to, oh, okay. >> I'm going to notice something here, which is that you, as you astutely pointed out earlier, we have two functions. Of P or to equations of P, and by the way, do you know what they are? [They're lines](#). >> Sure, because it's just a, it's linear in P, so that's what linear means. >> Right, so [what would happen you, do you think if we were to actually draw these lines?](#) >> I think we'd have two lines. >> Yes, and what would those two lines look like? Let's take a look, shall we? >> Sure.

$$15p - 5 = -10p + 5$$



Mini-Poker

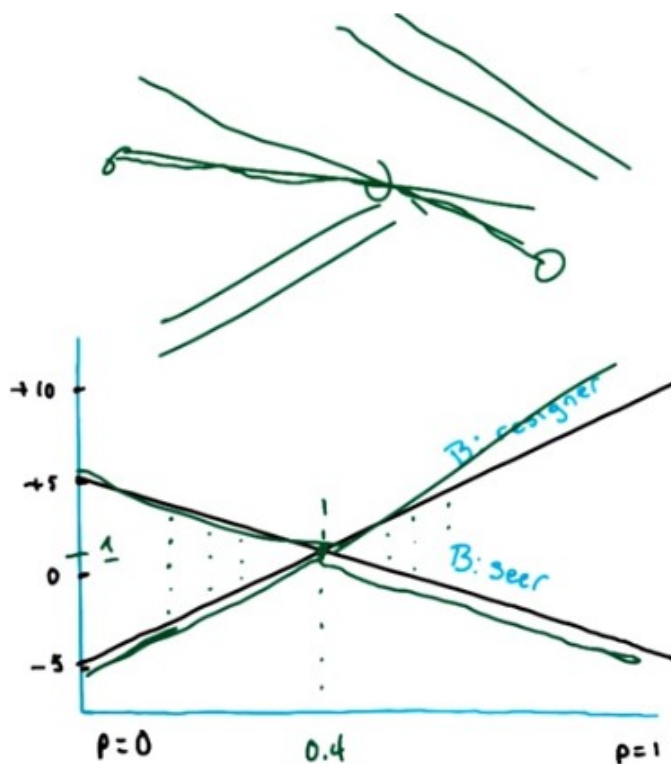
	B:	
	resigner	seer
A: resigner	-5	+5
holder	+10	-5

Quiz
Where do the lines intersect?

$$p = .4$$

19. Okay, so what I've done here Michael, is I have drawn both of these lines. So here's the line $15p - 5$. This is a case when b is a resigner. This is my best attempt at drawing it to scale. You start out minus 5, as you point out, and you end up with plus 10. And this is the line where b is a seer. That's minus ten b plus 5, so you start out with plus 5. And you end at minus 5. Okay? >> Cool. >> So, what do you notice about these two lines? >> They make an x? So they intersect. >> They do intersect. Can you tell where they intersect? >> How would I solve that? >> You'd have a quiz. Okay Michael, so here's a quiz. Where do the two lines intersect? Think you know how to figure it out? >> Yep. >> Okay. Then go.

20. >> Okay Michael, what's the answer? >> I don't know, but I would, here's how I'd get it. I'd set the equations of the two lines equal and find the p where the, where those values are equal. >> Oh, okay. So here let's do that. >> And I feel like that's like 25 p on one side and ten on the other >> Okay. >> So like 10 over 25, which is like 2 over 5, which is, like, 2 over 5. >> Yes, which is also 0.4. So 0.4, that's exactly right, and it's, you should do it exactly the way you said. Set the two equations equal to one another, do simple algebra, and. >> So what would have happened if p ended up being, like, not a probability? >> Then they wouldn't have crossed inside. >> Good point.



MINI-POKER

		B:	
		resigner	seer
A:	resigner	-5	+5
	holder	+10	-5

Quiz

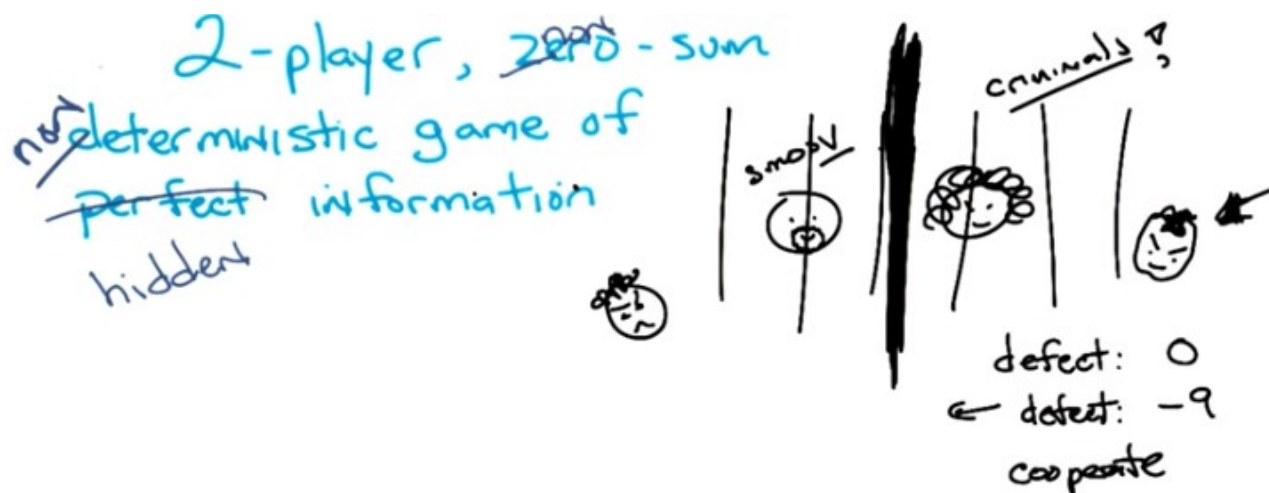
Where do the lines intersect?

$p = .4$

21. So by the way, here's a question for you. We can make it a quiz but I'm not going to make it a quiz. What is the value of the game at $p = 0.4$? >> So I just plug it into those equations, and so negative 10 times 0.4 is like, negative 4 plus 5, is 1. So I'm going to say 1. >> Mm-hm. So the value here is in fact \$0.01. And if you put it in the other equation you get the same thing. Yes, right, because they're equal there. >> That's actually the answer to the entire kit and kaboodle. If. >> Sorry, where were the kittens? >> If A chooses a mixed strategy, where with probability 0.4 he chooses to be a holder. Notice that it doesn't matter whether B is a resigner or B is a seer. You will end up here and there will be a value of +1 penny to A. >> On average. >> The expected value is. >> Yeah. >> plus 1. >> Neat! >> Now you might ask yourself, well what if B decides to be a smarty pants and also do a mixed strategy? What do you think would happen in that case? So of course it changes something, but it doesn't change the value of the game, because B, if B is a resigner, A is getting one on average. If B is a seer then A is getting one on average, and an, any average, any convex average of one and one is going to give us one. >> That's exactly right. So in fact if you think about it, if B tries to do a mixed strategy between these two lines. It's going to have to be for every single point between the two lines, somewhere the average is going to have to be somewhere in between there. No matter how you weight that average. So it's like a bow tie is this space of possible payoffs. Right and since for any let's if I just, If B decided to pick some, you know, some values such that it's in this region of the space. And it's going to end up somewhere, say between here that's fine, or here that's fine, or here that's fine, or here that's fine, or here that's fine. More importantly, right here it's going to have to be between the two lines. Where those two lines cross. So, no matter what mixed strategy B chooses. On average we will end up here, so the value of this game is the expected value of this game and it is, plus 1 for A. Okay so hang on why is that not there are other values that can be obtained. There is a minus 5 a plus 5, true. Why is it plus 1? Well its plus 1 here on average, the expected value of this game. Is plus 1. >> So you say that the strategy for A is to choose 0.4. >> Mm-hm. >> But why? Like, why is that, of all the different values. What [INAUDIBLE], just because the lines intersect. I don't understand what makes that a good idea for A. Like, it's not like it gets any additional payoff for having things intersect. Well, if a is going to choose a mixed strategy, and b is going to choose a mixed strategy. This is the mixed strategy for a, that

guarantees an expected value of $+1$. Meanwhile, let's imagine this is the way you're going to set it up. So we've been kind of talking in general that a's going to choose the strategy; b's going to choose a strategy, and they're going to go. And that, because we know everyone's rational, you already know what strategy b's going to pick, and b already knows what strategy a's going to pick. Right, because we made this assumption about everyone is trying to maximize their own utility, their own reward. So in a mixed strategy, if you know you're going to go for a mixed strategy, you have exactly the same situation. B can figure out well what is it that A should choose and A can figure out what is it that B should choose. So notice that in this particular version of the game, the way it's setup. Even if A announced beforehand to B. Listen I am going to choose to be a holder with probability 0.4. It doesn't matter what B chooses the expected value is this $+1$. Right? >> Yeah. >> But imagine if A said I am going to pick, I'm going to choose to be a holder with probability 1. Well then, what should b do? B should choose to seer. >> Right. >> It's exactly the situation we were in before. >> Okay, but here's the thing I'm having trouble wrapping my head around. So, it's not special that it's an intersection, what's special maybe, is it that, you know, if b is always. Okay; taking what you said before, that a is going to announce a strategy. So for anything that A can announce B is going to presumably do what's best for B which is to just minimize right? >> Mm-hm. >> So if you look at that triangle at the bottom. >> Mm-hm. >> If you think about that as those lines that V, upside down V shape. Yeah exactly that thing. If you think about that as being the payoffs, the payoff(收入) function for A, as a function of the probability P that it chooses to hold. >> Mm-hm. >> Then we've chosen the maximum. >> Right. >> Right, we're trying to find the probability for which A gets the highest expected value. And it's at that, it's at that peak there. >> Hm-hm. But notice something, Michael. For any case where the two lines cross, you're going to have a function of this form. >> Well, let's be sure of that, because that's, I wasn't seeing, I was thinking that that was not true. because basically in this case, we have a line of positive slope and a line of negative slope. >> Mm-hm. >> Right? So, what if we can have an intersection between two lines of negative slope. huh. So again, by doing the same exercise you were doing before, where you draw the, you sort of take the minimum of the two lines at all points. >> Yeah. >> You'll end up with this. Where do you pick it then? >> The far left, $p = 0$. >> Yeah. >> But not the intersection in particular. >> No, that's true. >> Okay, alright, I just, wasn't quite getting that. But so, the intersection is special, because in some cases that is where the max is. It can only I guess the max can only be in those three places, right, it could be far left far right or the intersection. There's no other way to have a maximum of lines. >> Right. And b by the way there's another case like this where they never actually cross. >> Mm. >> Or like this where they never actually cross. Got it. >> Right, but in either case, right, it's always going to be those three points. It's going to be the extreme or where they cross. If they happen to never cross, then that point goes away and you just pick the maximum. So what you could, so in fact, if you wanted to think of an algorithm to choose the right thing to do for A, you basically plot the lines. You take the min at every point and you find the maximum point. >> Oh I see, so you could just kind of discretized it almost. >> Yeah, and you're done. Seem reasonable? >> Well discretized seems problematic but like I get that we're trying to find the probability so that we maximize. Oh wait we maximize the minimum of the two other things. So it's like maximin again. Yeah. >> Except here, so, in fact, it is exactly min and max or max and min, except, oh, I guess, if you're thinking from A's point of view. It's min and max or max and min but in this case, there's this other parameter, which is the probability. Which is how you determine how you're doing the min and max. >> Got it, right. In this bigger space, it's min and max. I see, yeah, yeah, yeah, that makes sense. What about so, why is A the one who needs to be random? Why isn't B the one who needs to be random? >> It's, you end up with the same, you end up in the same place. Because again, remember, you're, you're doing this kind of optimal notion, where underneath all of that is this belief that both people are going to be rational. So the only reason to do this, to take the, you know, maximum minimum, is if you believe that B is always going to try to minimize what you do. But B's in exactly the same situation. Right, from B's point of view, That's what's happening (即左下角圖中的 max 組成的,

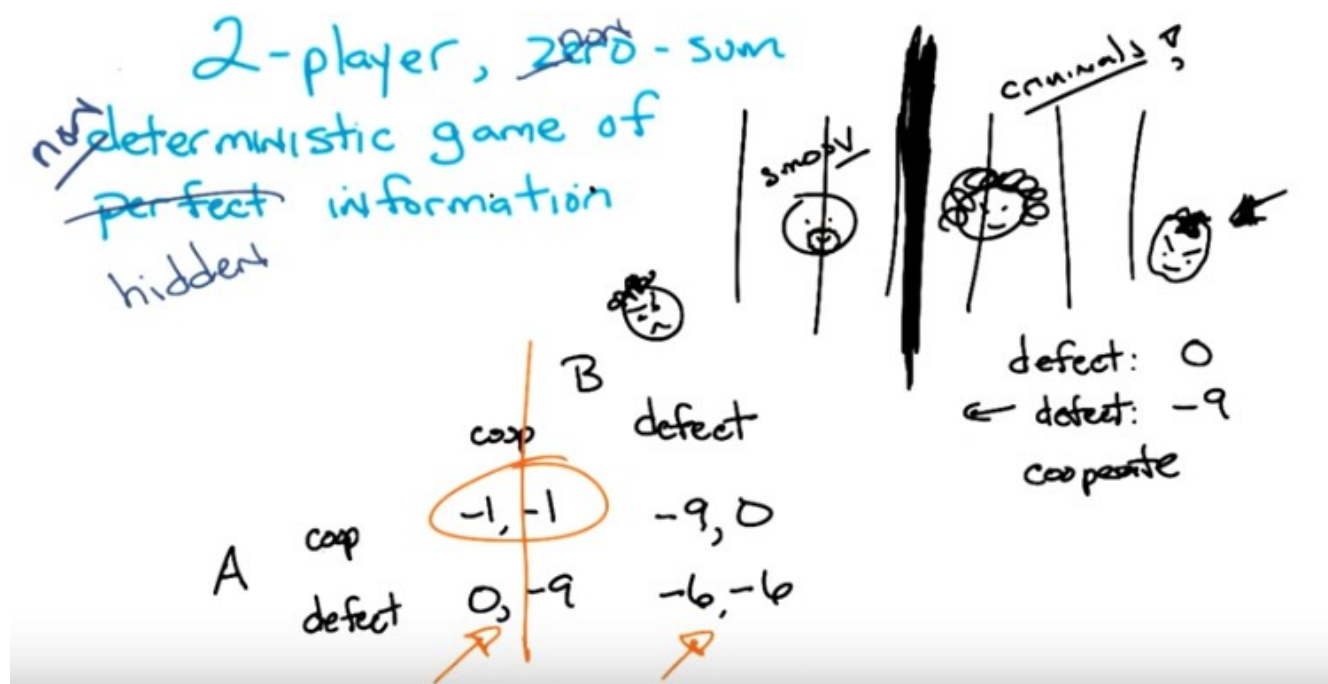
即那個 V 形的). >> Huh. It's the same equations? >> Yeah, it just it gets turned around. That's actually an interesting exercise. Maybe that's the next homework assignment. >> Okay. >> So, did that all make sense to you? >> [LAUGH] [CROSSTALK] It looks a little green and scribbly, but yeah. Sure. That's cool. >> That's how you know when you're done, Michael. When it's all green and scribbly. >> [LAUGH] Does this generalize to more than two options? It seems like it's going to get messy fast. >> Yes, it does. It generalizes to more than two options. Effectively now, you're just doing a max over n things, instead of two things. >> And you have to search for, there's possibly way more intersections to worry about. >> Yeah, but all you care about is the minimum. Think of it as you're always drawing the minimum function.



上圖的意思, 看下圖就明白了

22. Okay Michael, so, feels like we've relaxed almost everything we could relax except for one thing. We're now going to look at two player non zero sum. None possibly none deterministic games of hidden information. >> Cool. >> And that's going to turn out to be messy, but get us to the place where I've secretly been trying to get us all along. So, let me describe a game for you. Very carefully, and let's see where it leads us. Okay, here's the game, you've got two people. These two people are criminals. >> Oh no! Are they smooth criminals (犯罪高手)? >> One of them is. >> [LAUGH] [MUSIC] >> [LAUGH] Oh, you're terrible. Okay, so we have two people. They're criminals. Okay, and unfortunately they've both been captured by the cops. I have actually no idea how to draw that, but let's just try to draw that like this. The cops come along, and capture them both because they are suspected in a particular robbery. Okay? >> Hm. >> And they take them and put them both in jail. So those are jail bars. >> [LAUGH] >> Okay? But they don't just put them in jails. They put them in two separate jails. >> Oh no. >> Okay. And one cop goes to one criminal and says listen, here's the deal. We know you did it. Okay? We know you did. And your friend over there, he's currently singing, singing like a bird. And he, is going to pin it all on you. So this is your last chance to give us a little help and admit that you two did it. Or, that the other guy did it. You admit that the other guy does it, you say it was his fault, then we'll cut you a deal. Now to make it worse, the cop tells him that there's another cop over there. >> [LAUGH] >> Who's talking to the smooth criminal, and is offering him the same deal. >> Hm. >> And whoever, goes first, whoever pins it on the other guy first, gets to walk. Okay? >> Hm. >> So, if I can get the curly-head guy to defect (叛變), okay? >> Hm. >> Then, I am going to let him walk and he will spend zero months in jail. >> Okay. >> Okay. On the other hand, if the other guy defects, then he's going to get to spend zero time in jail. And since we've got all that we need. We got a confession. To get this guy to go to jail. He's going to go to jail for nine months. Okay? >> Negative

nine months? >> Yeah. Oh, that's a cost in months. I see. Uh-huh. [CROSSTALK] >> Yeah that's a cost in months. Okay? So if, curly-head guy defects before the other guy does, then zero. He walks. He pays no cost, other than the cost he's already paid for a life of crime. >> [LAUGH] >> Now if he refuses to drop a dime on the other guy, but the smooth criminal decides to defect he's going to lose nine months. That make sense? And the other guy's going to walk. So he's getting the same deal, they're both getting the same deal. You got it? >> Yeah. >> All right, now. It's a little more complicated than that, all right? But I, I hope that you're getting all, keeping all of this in your head. It's a little more complicated than that. There's actually four choices here, not just two. It's not just a case of. I defect or the other guy defects. I defect, the other guy doesn't. I don't defect, the other guy does. Okay? You could both refuse to drop a dime on the other. You could cooperate, or you could both rat out (詛咒, 責罵) the other. >> At the same exact moment? >> At the same exact moment. >> So, there's a big thick wall here. Curly guy doesn't know what smooth guy is doing, and smooth guy doesn't know what curly guy is doing. >> Right. >> So the question is, if you have a choice between either defecting or. >> Which means blaming the other guy? Which means blaming the other guy you defecting from your friendship. >> Oh. Or you can cooperate, that is be true to your friendship. I can, we both people can defect, both people can cooperate, one person can defect and the other person cooperate, so there's actually four different options there. >> I fee, I feel a matrix coming on. >> Yeah, I think there's a matrix coming on. So let's draw the matrix, because I think that makes it easier to see. But you have the background here, right? >> I think so, yeah. >> The key piece here is that each person has a choice of either defecting from their friendship or trying to cooperate. Keeping their mouth shut, and, they don't know what the other person is doing. So, what are all the costs here? What are the worst case things? Let's just draw it all out as a matrix.



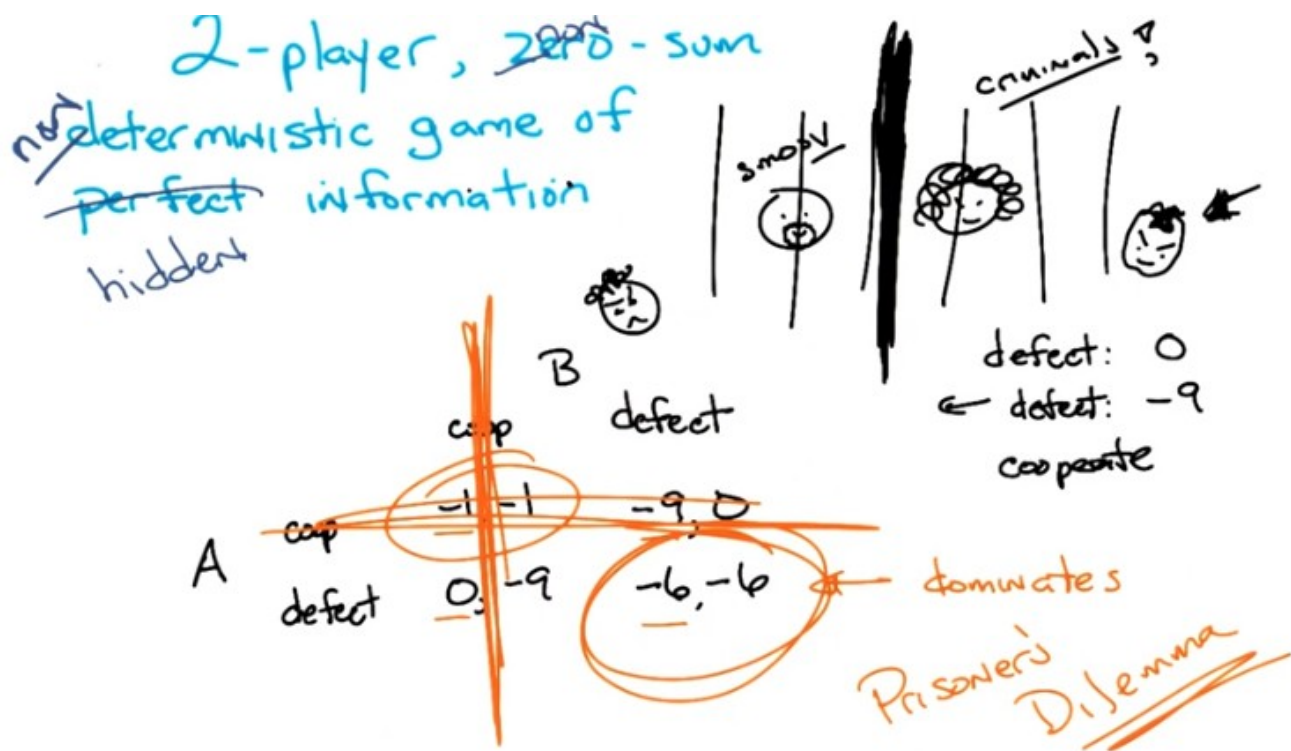
defect 一詞的意思是叛變, 即出賣自己的隊友
cooperate 即保持沉默, 即跟自己的隊友 cooperate

上圖的意思是, 有兩個犯人, 被隔離審問, 每個人都可以選擇 defect 或 cooperate. 上圖 matrix 中的元素的意思是, 比如(-9, 0), 表示 A 坐牢 9 個月, B 坐牢 0 個月. Matrix 的意思就是: 若兩個都 coop, 都只坐牢 1 個月.

若有一個人 defect, 另一個人 coop, 則 defect 那人坐牢 0 個月, coop 那人坐牢 9 個月
若兩個人都 defect, 則兩人都坐牢 6 個月。

23. So let's just call the smooth guy A, because that's what we've been doing all along. And he can either choose to cooperate, or he can choose to defect. Now by the way, I'm saying cooperate and defect here because they're terms of art, not because these are the words I would have chosen. Okay? B can do the same thing. B can choose to cooperate or B can choose to defect. Now I've set up the game in a particular way here. The, the, the cops have set up the game in a particular way here. If B defects but A cooperates, then B gets to walk and A has to pay the price, nine months in jail. So, I'm going to put into this cell minus 9, 0. So, this means this is the value of this set of strategies for A, and this is the value for B. So the first number. In the pair is what A gets and the second number's what B gets. Okay? >> Got it. >> You with me? >> Yep. >> Okay, now notice we have to do this now because **it's no longer zero sum. It's not going to always add up to a constant.** >> Oh. Well, it is, so far it's a constant negative 9. >> That's right, oh and in fact it looks this way as well because I have a symmetric deal here. So if A defects and B cooperates, A gets to walk and B goes to jail for nine months. Okay? >> Yep. >> So right now, you're right, it's looking like a zero sum game, but it isn't. Because what happens if both A and B drop a dime on each other? Well, they both confessed. So it's a little good, it's a little bit better than having one of them, only one of them, confess. The deal that the DA is willing to give them is that both of them will spend six months in jail. Now these are just, these are just, the, the numbers that are a part of the game. This is, I'm not computing this from anywhere. This is just the way the, the cops have set this up. And I'm going to, now what's an interesting question here is what happens if both A and B keep their mouth shut? They both choose to cooperate? It turns out that it's not a perfect world because they were caught with a bunch of guns they didn't have permits for. So if neither one of them admits to robbing the bank, they're still going to do a little bit of time. But in this case, it's a small weapon's charge and so each only spends a month in jail. >> I see, so now it's definitely not zero sum. >> Mm-hm. >> Because we have a negative 2 there, a negative 12 there, a negative 9 and a negative 9. >> Right. >> Okay. >> Okay. >> So, you got the game, you understand it? >> Yeah, I think so. >> It's very simple. Now, just looking at this, **what's the best possible outcome for the duo?** >> So if they cooperate with each other, then there's you know, one month later, they're back on the streets, back to their criminal activities. >> Or they're reformed. You never know. >> Sure. Back to their choice of whether to have criminal activities. The, the mutual defection one, that, where they both defect, it's either, well, 12 months or six months, depending on how you think about it, but they don't do nearly as well. And then the defect and cooperate, it seems like there's a lot of incarceration that will happen. >> Mm-hm. >> So, **it feels like the best for the, for the two of them is to mutual cooperate.** Cooperate, cooperate. >> Yeah. And that makes sense. This is sort of what you want to happen. Both of them keep their mouth shut. They do a little bit of time, but on average, they do pretty well. Right? >> Yep. >> **So my question to you is, is that going to happen?** >> Sure, **why wouldn't it happen?** >> Well, you tell me. **If I know that you're going to cooperate.** Let's say that I'm A, okay? And you're B, and I know that you're going to cooperate. **What should I do?** >> **You should cooperate.** >> **Should I?** You've chosen, you've chosen this column. All right? >> I see. Well, if you think of it that way then it's not a joint choice, but it's actually an individual choosing. **Then you're better off defecting because then you get off scot-free.** >> Yep. >> Then, but I go to jail for nine months. I could have a whole baby in that time. >> Usually takes closer to ten but sure. But you go to jail for nine months, I don't. >> So if you are just that cold. I guess because you're a criminal. >> It doesn't matter. Remember, the matrix is everything. The value of doing this to me is 0 versus minus 1. >> Alright, that's, that's cold, man. >> I agree, but it's what the numbers tell you. What if it wasn't criminal? What if it was just, you know, the amount of money that I was going to win in a, in a mini poker game? >> Yeah, but it's a different kind of mini poker game, because we're both, because. [LAUGH] It's like you beat me. But, by beating me that way, you like, kill, nearly kill me. >> So? Are

you saying that when? >> What do you mean so? >> Are you saying you always let me win whenever we play poker? >> No. >> Okay then. Because you're cold, is that what you're saying? >> No. Wait, what? >> Exactly, right. So the point, Michael, is that if I know you're going to cooperate, you're going to choose this column, then I should defect. >> Okay, alright, so, fine. So now I'm going to jail for nine months. >> Mm-hm, if you choose to cooperate. >> Aha. If, if is good. So what you're saying is, I could drop a dime on you. [LAUGH] >> You could. So, you could choose to defect, and if I knew you were going to choose to defect, what would I do? >> Well, you already, you already showed your colors, man. You're, you're defecting, so I'm just switching, I'm just saving myself three months by ratting you out. >> Yeah, so we would end up here (both 6 months in jail). By the way, you know that this game is symmetric, right? >> no. You defected first. [LAUGH] >> No, here's the thing. Since the only thing I care about is maximizing my own reward, my own value. I'm going to do the thing that makes sense for me to do in this case which is if you cooperate, defect. If you defect, defect. But you would do the same thing because your whole goal here is to maximize your value. >> Yeah, well you don't know me like that. >> Yes I do because it's everything that's here in the matrix. This is the wonderful thing about game theory. All the stuff that you're concerned about is all inside the matrix. Remember the rules of the game don't matter. There's no prisoners here. There are no criminals. There's just, I get a dollar, I lose a dollar, or I lose zero dollars. I lose \$9 or I lose \$6. Or, for that matter, cents. I lose one penny or I lose no pennies. I lose nine pennies or I lose six pennies. It doesn't matter. This is the value of these particular strategies. And I'm going to want to defect if you cooperate. You'll notice if you defect, I'm also going to want to defect. In fact, if you look closely at this matrix, you should notice something, Michael. From A's point of view, when does it make sense for me to cooperate versus defect? >> You mean if you have a, some kind of probabilistic policy over cooperate and defect? >> No, no, not even that. Just in general, is there ever a time as A that I would rather cooperate than defect or vice versa? >> Oh, I see. So if I know you're going to cooperate, I should defect to, to save myself a month. But if I know you're going to defect, I should defect to save myself three months. Either way, I'm coming out ahead. >> Right. So in fact, let's take a look at this.



24. If I look at the value for me as A, between cooperating and defecting in this first column. It's minus one versus zero, right? Which number is bigger? >> [LAUGH], zero. >> Right. >> If I look at it in this column, it's minus 9 or minus 6. Which number's bigger? >> Negative six. >> Right. In both cases, defecting is better than cooperating. So in fact this choice, this strategy, dominates the other. In other words it is always better for me to defect than cooperate. So I will never cooperate, ever. Because it's always better for me to defect. >> So fine. So I'll never cooperate ever. That'll show you. >> By exactly the same argument, minus one versus zero, minus nine versus minus six. Defecting is always better. B will never cooperate. What's the only thing that's left? >> Pain. >> Pain. This is called the Prisoner's Dilemma. >> Huh. You said it was simple, but it seems kind of evil. >> I didn't say it wasn't evil. Those things aren't opposite of one another. It's not like perfect versus hidden. Or pure versus mixed. It's not easy versus evil [LAUGH]. Evil's often actually easier. >> It's easy and evil. >> So we're in a little depressing place here, Michael. You claimed in the beginning, and I agreed with you, that this (都 coop) is sort of the best option, you want everyone to cooperate, because that's sort of what's best for the group. But because defecting dominates cooperating both for A and B you're going to end up here (都 defect). At least if you reason about it that way. >> Yeah, I see that. >> Hence prisoner's dilemma. The only way to break this is if somehow we could communicate and collude with one another. So that we could guarantee that we would both choose to cooperate at the same time. >> So that wall? What about that wall? What if you put a like a Skype connection or a Google hangout? >> Well, if you were forced to say what you were going to say at the same time, or somehow be able to punish someone maybe for, you know not doing the right thing, then you might be able to make a different decision. But for this very simple version where I'm going to do it once, even if I could hear what you had to say, if one of us, whichever one of us went first, the second one would always be able to take advantage of that. >> I do find this kind of depressing. >> Yes, it's a dilemma. It's a true dilemma. So, this brings us to a more general sort of strategy. This whole notion of strict dominance, works in this case. But, you could imagine very complicated large matrices where it may not work. But it turns out there's a generalization of this notion of dominance. That works remarkably well. And that's what people tend to use to try to solve these kinds of games, to find out what the true value of a game is. So let me describe that for you, ok? >> Okay.

A BEAUTIFUL EQUILIBRIUM

n players with strategies

$s_1, s_2 \dots s_n$

pure
&
mixed!

$s_1^* \in s_1, s_2^* \in s_2 \dots s_n^* \in s_n$

are a NASH EQUILIBRIUM iff

$$\forall_i s_i^* = \operatorname{argmax}_{s_i} \operatorname{utility}_i(s_1^* \dots s_i \dots s_n^*)$$

no reason for any one to change

25. Alright Michael. I'm going to define a new concept for you. It is called the Nash Equilibrium. >> Nice. >> Okay, here's the set up. You have n players. So we move beyond simply two players. You have n players. And each player has strategies that it can chose from. So here I'm referring to them as, s_1 , these are all the strategies for player 1. s_2 are all the strategies for player 2 up to s_n all the players for strategy N. Got it? >> Yep. Okay so each of these are set, so I'm going to say that the particular strategies, s_1^*, s_2^*, s_3^* , so on and so forth s_n^* that is a strategy that each of the N players has chosen, is in a Nash equilibrium if and only if for each one of those strategies chosen by the n players it is the strategy that maximizes the utility for that particular player, it is the strategy that maximize the utility for that particular player given all the other strategies that were chosen. Now I actually find that difficult to kind of work through, so let me try to say it a different way. Given that you have a set of strategies $s_1^*, s_2^*, s_3^*, \dots, s_n^*$ We know that they are in Nash equilibrium, if and only if you randomly chose one of the players, and gave them the chance to switch their strategy, they would have no reason to do it. >> Interesting, okay. >> So does that make sense? >> Yeah. >> Right. So an equilibrium right, just in general, the word, the word sort of makes sense, right, an equilibrium is at a place where everything is balanced. And in some sense, there's no reason for anything to move because they're in balance. So we set that a set of strategies are a Nash Equilibrium if no one person has any reason to change their strategy in a world where everyone else's strategy remains the same. Then you're, then you're in equilibrium, and in particular you're in a Nash equilibrium. >> Nash equilibrium. Okay good. >> Do you know who Nash is? >> Ogden Nash was a poet. >> Not that Nash. >> I think there was a TV series with a Nash in it. >> Yes there was. >> But you're thinking of John Nash. >> That's right The Nobel Prize winning person who was the, was featured in the movie A Beautiful Mind. >> Yep, that's exactly right. And it was, in fact, a Beautiful Equilibrium. Okay, so, there's the notion of a Nash Equilibrium. It was admitted by John Nash, if you believe the movie, in order to pick up women. Which I, you know, you know I never, I've never seen this movie. >> Oh really? Oh, I watched it and it was, it was surprising how unhelpful it was in explaining what a Nash Equilibrium was. >> [LAUGH] I'm not surprised. I hope this is helpful, though. So, it really is a kind of difficult

concept to completely wrap your, your mind around, but what it really boils down to is, listen, if we all picked a bunch of strategies and we knew that one other person, one of us, we don't know who before hand, but we know that one of us Would have the opportunity to change their strategy after they see what everyone else's strategy is. We'd be in a nash equilibrium only if that person, whoever that person is, has no reason to change their strategy. >> Gotcha. So, so specifically you've made a distinction between strategies that were pure and strategies that were mixed. >> Right. >> And I guess I don't see in this case which kind we're talking about. >> Right. So, in this particular case, [we're talking about pure strategies, however, exactly this wording works with mixed strategies.](#) >> Oh, I see. So you could have a pure Nash equilibrium or a mixed Nash equilibrium. >> Right. And so, [you could, each one of these, instead of choosing a particular strategy and saying they're a Nash equilibrium, you could talk about a probability distribution over each of these sets of strategies and say those are Nash equilibrium if. No one would want to change their probability distribution.](#) >> Got it. >> So this works for both pure and mixed. Alright, so you think you understand this? >> [LAUGH] sure. >> Good, we will test that.

26. >> Okay, Michael. So you think you know everything, so here's a quiz. So, [here are two matrices, matrix prisoner's dilemma and matrix bunch of numbers that look vaguely symmetric but aren't quite.](#) So here's what I'm going to ask you do do. [In each of these cases find the Nash equilibrium.](#) >> Rut-roh. So do, you told me what it was but you didn't tell me how to find them so that seems kind of rude. >> It was implicit. It's intuitively obvious even to the most casual observer. >> Oh, well then, okay. So, and just to remind me now, [each row and column is a choice for one of the players or the other player.](#) >> Yeah. >> [And the first number in the pair is the row player or A's pay off, and the second number is B's pay off.](#) >> Yes. >> And we might need probability distributions, or we might not, depending on what it takes to be a Nash equilibrium. >> Right. >> okay, I don't [LAUGH] I'm not, it's not like I see the answer yet but it, but I either will be able to find it, or I won't. At least I understand the question. >> Okay. And by the way, just to make it easier for you [they are pure Nash equilibrium,](#) okay? >> Hm. >> Alright, so you ready? >> Yep. >> Just circle the one or underline it or whatever it is Pushcar does to make it so that you can answer this quiz. You ready? >> Totally. >> Go.

A BEAUTIFUL EQUILIBRIUM

FIND THE ^{pure} N.E.:

B

	-1, -1	4, 0
A	0, -9	-6, -6

(strictly dominated)

QUIZ

B

	0, 4	4, 0	5, 3
A	4, 0	0, 4	5, 3
	3, 5	3, 5	6, 6

27. Alright Michael, you got the answer? >> Yea, I'm ready to try and figure it out. >> Alright let's go. Let's try the first one. >> I feel like the first one's going to go rather well. So we need a pair of strategies so that no player is. Motivated to switch. >> Mm-hm. >> And you told us they were pure strategies, so we actually have a nice algorithm for doing this, which is we could just check one of them with the definition. >> Mm-hm. >> But, but in the case of prisoner's dilemma, I think a natural place to start would be that (-6, -6). >> Mm-hm. >> So let's say that A chooses the second row and B chooses the second column, let's see if that's a Nash Equilibrium. So both players need to be happy. So if, would A be happier switching? If A switched, it would be getting -9, which is worse. >> Mm-hm. >> So A is happy, where A is. And if B switches, B would be getting -9, so boom, Nash Equilibrium. (為何不變到(-1, -1)去? 因為 Nash equilibrium 要求只能一個人變, 其它人不能變) >> Done. >> Now I didn't verify that other ones weren't a Nash Equilibrium, but you didn't say. Find all of the Nash Equilibrium. You just said find, well you did say find the. So you kind of implied that there's just one. >> Right. But actually that's true, but you don't have to check this because we already went through an exercise where we, where we knew the answer was minus 6, minus 6. And the way we did that was we noticed that for A defecting the second row is always better than this. Right? And then in particular this (second) row strictly dominates this (first) row. >> Right. >> Which implies that if I picked anything on this (first) row I would rather move to the other (second) row. And you can see it. -1, 0 is better (than -1). -9, -6 is better (than -9). That's what it means to be strictly dominated. So, I'd never pick this (first) row anyway. And this same argument for B for this column. So, you'll notice that by getting rid of the things that are strictly dominated, the only thing we're left with is this. And, it turns out, in fact, to be a Nash Equilibrium. So, this is correct. >> So, you just told me how to do my job, which makes me little sad. >> Well, in this instance. >> Because I already had the answer to this. But maybe, maybe you were signalling to me how I might attack this next problem. Maybe. >> Maybe. >> So, A is choosing a row and gets the first number. So, is there a row, that, were they dominates all the other rows? It doesn't seem that way. >> Mh-hm >> But maybe, maybe, oh I could start with B, maybe B, there's a column that dominates all the other columns. But no, it looks like it's totally symmetrical. >> Yep, so strictly dominated doesn't necessarily help here, and by necessarily I mean doesn't. >> So that was kind of mean. Thank you. >> All right, so oh, but we have something else we could do. >> Yes. >>

So, there is a, the largest number that anybody can get is 6. >> Mm-hm. >> And there's a play where both of them can get the 6. >> Yeah. >> So there's no way they're going to want to switch away from that, because everyone's getting there kind of maximum out of reward. So A bottom row, B right column, gets us Nash Equilibrium. >> And that is in fact correct, and you can see it because from here I would always, it would always be worse for me, and it would always be worse for me. So, these are in fact the Nash Equilibrium, for these two problems. >> Cool. They've seemed easier than I was expecting. >> Mm-hm.

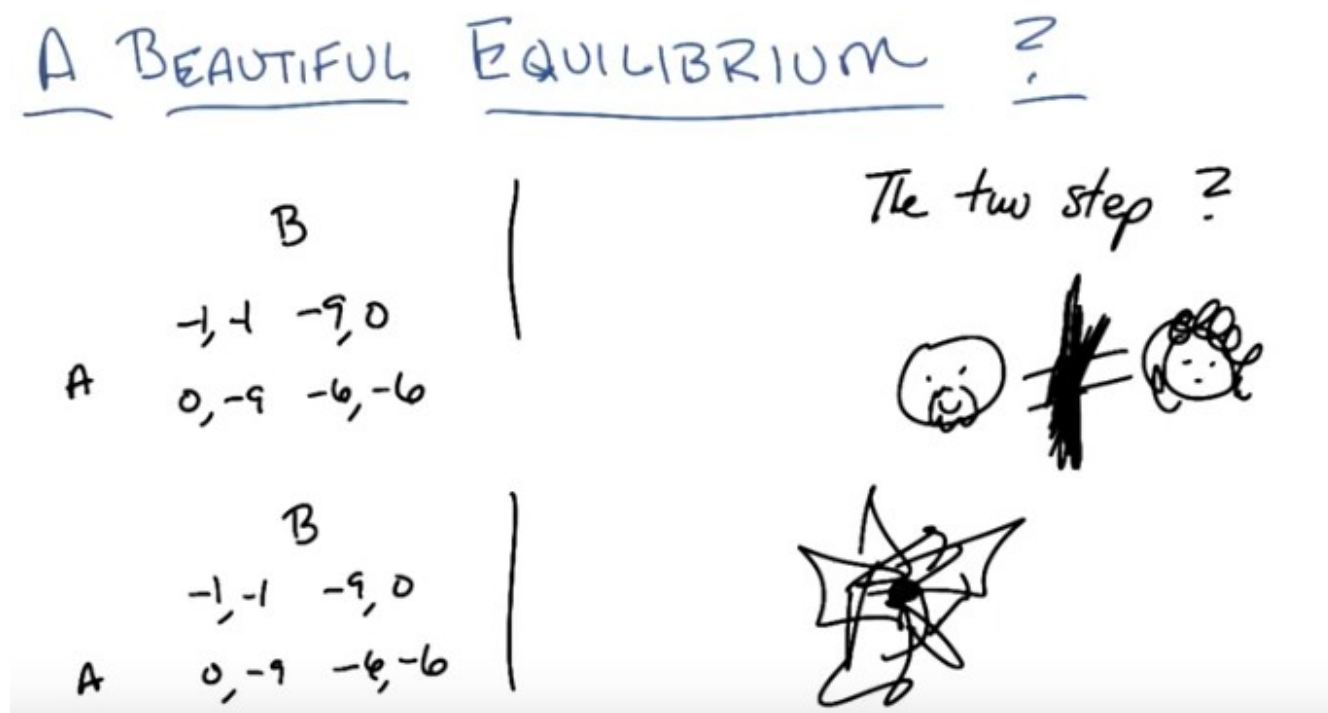
A BEAUTIFUL EQUILIBRIUM

- In the n -player pure strategy game, if elimination of strictly dominated strategies eliminates all but one combination, that combination is the unique N.E.
- Any N.E. will survive elimination of strictly dominated strategies
- If n is finite & $\forall i$ S_i is finite \exists (mixed) N.E.

注意第一條中是 dominated, 而不是 dominating.
 右下角括號中是 mixed

28. So here are three, fundamental theorems that come out of all of this work on Nash equilibrium. They're all in the screen in front of you. I'll read them for you in case you can't read my handwriting or if there's some typo you discover that forces me to erase some words and then write some new ones. Okay, are we ready? Alright, the first one is, in the n -player pure strategy game, if elimination of all strictly dominated strategies, eliminates all but one combination of strategies, then that combination is in fact the unique Nash equilibrium. So that's, that's what happened in prisoner's dilemma, we got rid of all but one option and that option had to be the unique Nash equilibrium. >> You say, eliminate all of them, do we, is it possible that things that we couldn't eliminate in one round, we could eliminate in the next round? >> Yeah it's possible, so in fact, this is done in an iterated fashion. You get rid of whatever you can eliminate and pretend they were never there. Because the truth is, no one would ever choose them. And then with what you have left, you just do it again. So it is possible. Although not in any examples that we did. >> Okay. >> The second one is and this, both of these I think are kind of Obvious. They sort of make sense anyway. At least to me. That any Nash equilibrium will survive the iterated elimination of strictly dominated strategies. In other words if you get rid of things that are strictly dominated you will not accidentally get rid of Nash equilibria in the process. And that makes sense because, if they're strictly dominated then if you ever end up there, you would want to leave. And therefore can't be a Nash. >> And therefore can't be a Nash equilibrium, that's right. So, those two

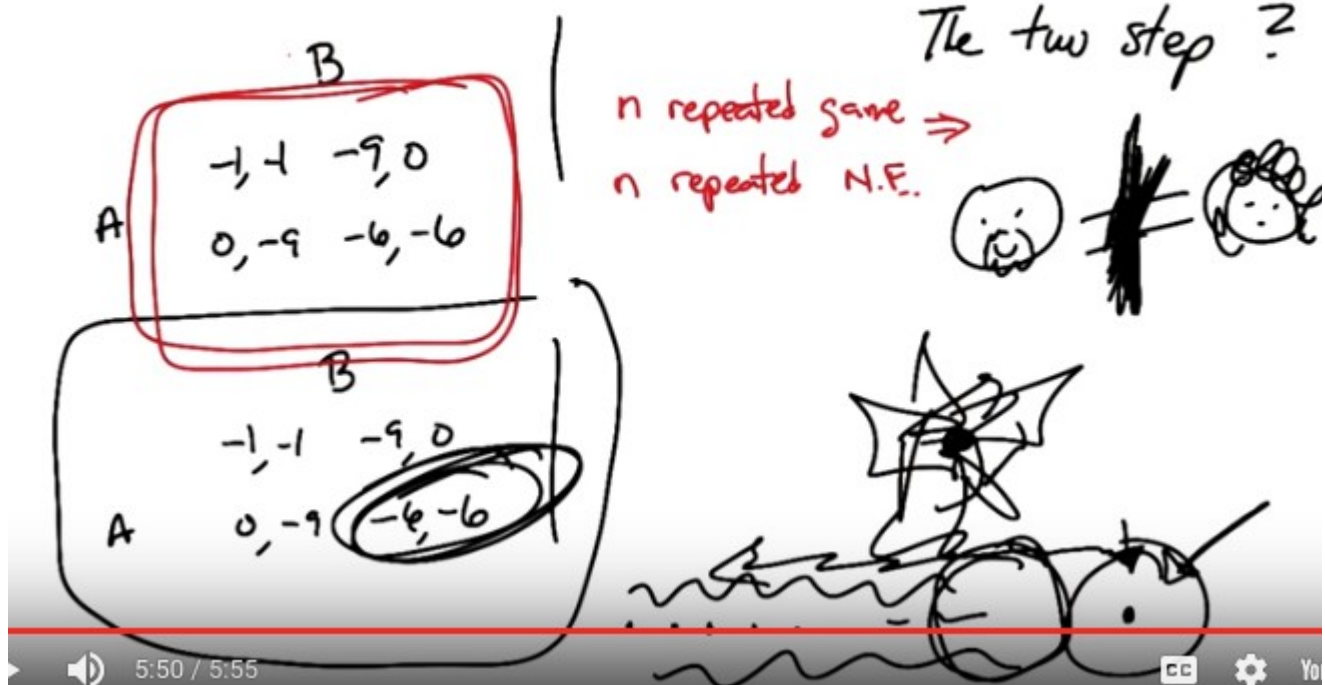
things sort of make sense I think. The last one is true but isn't at least painfully obvious, at least not to me, and that is. >> If n is finite, that is you have a finite number of players, and for each of the set of strategies, that set of strategies is also finite, in other words you're still in a finite game, then there exists at least one Nash equilibrium which might involve mixed strategies. **So you're saying there's always a Nash equilibrium?** >> **Yes, possibly mixed, for any finite gain.** So, when you said I relaxed everything, I didn't actually relax the requirement that it be a finite gain. >> Fair enough. But you stopped writing finite, so it's sort of the same thing. >> Yeah, I agree. Because why would you play an infinite game? >> I don't know. Maybe you got a lot of time on your hands. >> Mm. That would be like an infinite jest. [LAUGH] >> Okay. So those are the, the main, results here. So, what did we learn? What we've learned is that with interesting games. We end up in these weird situations where we have to figure out how to solve the game and one really nice concept, is this notion of an equilibrium and in particular the notion of a Nash equilibrium. >> Cool. >> And there you go.



29. Let me just wrap up by doing a couple quick things. **Earlier on when we were talking**, you made a kind of offhand comment about, or maybe I made an offhand comment, about the fact that **because we're doing this little prisoners dilemma thing, we have this sort of wall here, we're going to end up in this little bad place. Right? With the -6, -6.** And there was this kind of notion that what if we could hear each other, would that make a difference? And the answer was, no not really, because whoever goes first is going to lose. And if you try to get people to go at the same time, well in some sense, that's the same thing as having a wall and you can't hear each other if you have to go at exactly the same time. Right, does that make sense? >> I didn't quite understand. Why does the first, person who goes first lose? >> Because if I find out whether you're going to cooperate or defect, then I will just then do the thing that makes sense. So if you do anything but defect, I will defect, and then I will win and you will lose. So whoever goes first runs the risk of the other person screwing them over. So. >> I see. Unless, unless that person defects. >> Right, but if that person defects, the other person's going to defect. >> Right, so, but you don't lose any more than you would have. >> That's true, and, but you still end up in this kind of unfortunate place. So, one question you might have there is, yeah, but that's just because you can only do this once. But **what if you could do this twice? So, lets imagine were, going to be in**

this situation today and then were going to be in this situation tomorrow because, you know, you keep leading me astray(誤入歧途). So, whatever happens today I can use as information, for what I might do tomorrow. So, if I decide to go ahead and cooperate this time. And you defect, and I know you're the kind of person who defects, and so when we're in this situation again, I am definitely going to defect. So maybe in that situation, it's in your best interest to go ahead and cooperate because, then we can keep cooperating together, and in the end it sort of works out. So maybe this notion of not just playing prisoners' dilemma once, But playing prisoner's dilemma twice or three times, or four times, or five times, might come to a different result. >> Yeah, that seems plausible. I mean, I could imagine saying to you though this channel in the wall, something that says, you know, cooperate with me or, yeah, right, or I will, I will stop cooperating with you. And so, you get more reward if you, if you cooperate with me. >> Right. So, my question to you is, what does happen exactly in this set? Where I have two versions, two consecutive games of prisoner's dilemma to play. What am I going to do? >> Ok, well now I mean, we could turn this into a game, right? I mean, that's what you were teaching us how to do. >> Yeah. >> So you, so now the game has four, well, I don't know. So one way to do it is to say A has a choice to make on the first step, and [COUGH] that could be cooperate or defect, and then the second step could be cooperate or defect, so there's four possibilities. >> Right. >> B has the same four possibilities. But maybe we can actually have A be responsive to B. So A has two things to do on the first step and then two things to do on the second step for each thing that B did on the first step, for a total of eight combinations. >> Right. You could do that. [CROSSTALK] >> So if we made an eight by eight matrix, we should be able to solve it. >> Right. So, that's pretty easy to do. You just kind of draw an eight by eight matrix in, you know, that many dimensions. And it, you know, you end up right here. >> I wasn't thinking about that many dimensions. I just thought, like the normal kind of matrix. But, okay. >> Oh, I see, I see, I see. >> I mean its eight by eight, so that's kind of a pain. Right. So, its 64 cells and, I guess, I guess we don't want to fill that out. >> Yeah, but you know. We could fill that out. But, I'm going to help you out here by pointing out You know what? It's not going to make a difference. >> Oh. >> So, here. Let's see if I can walk you through why it's not going to make a difference.

A BEAUTIFUL EQUILIBRIUM ?



30. Let's imagine I have 20 of these games. Okay? >> Sure. >> So I'm going to play these games 20 times in a row. Now what you've been doing is you've been going forward. You've said, well, if I can get us the right thing or if I can basically threaten you and say. If you screw me over this time, I'll screw you over from now on. Then maybe I can get you to do the right thing and vice versa, since it's a symmetric game. And then we'll end up doing the right thing all along. And that makes a lot of sense, right? >> Yeah. >> If you're going forward. But what if you're going backward, Michael? Let's imagine we are doing [CROSSTALK]. What is amazing to me is that is exactly backwards what you just said. So, let's imagine we are doing these 20 dot dot dot dot, and now I'm on the 20th game, okay? What's going to happen on the 20th game? >> Well, I mean, I guess I could have built up trust in you from the previous games and think that you're going to cooperate with me, because we've been cooperating together. And then that would be the perfect time to drop a dime on you. >> Exactly. So, if this is the 20th game and we're in this situation, remember whatever value we've been adding up along the way, that's sunk cost. Right. The only thing that's left is the final game. And the final game looks like this, which means we're going to end up here (-6, -6). But guess what, Michael? The final one is determined. So since we already know the outcome of the final game, the only one that we, the next one that we can look at is the 19th game. Well, we already know what the outcome of the final game is. So this is effectively the last game. >> Oh. >> And what is the outcome of that going to be? It's going to be this and backwards, backwards, backwards, backwards proof by induction because the only that proves computer scientists know how to do is proof by induction. It turns out that we will always defect >> For what it's worth, those are the only proofs worth doing. But, okay. >> That's true. It's a fine point. Truth by induction and perhaps by conduction [LAUGH]. So, there you go Michael, even if I can play multiple games to try to build up trust, the truth is the Nash equilibrium, if I filled out that eight by eight matrix you wanted me to fill out. I would still end up in the same place where I would defect both times. >> Certainly if, yeah okay, so if you're going to be a jerk then, then I have to be a jerk, but yeah, I guess that's right. That's, that seems pretty horrible. >> It is. And by the way this is not just something

I'm making up. This is another theorem that comes out of Nash equilibrium. That if you have an n repeated game then the solution is n repeated Nash equilibrium. So whatever the Nash equilibrium is for the first game, the one version of the game, is the repeated Nash equilibrium for all versions of that game. >> Okay, wait a, hang on, hang on. So I. Okay, I buy that. >> Mm-hm. >> But couldn't you, well so, in a game that has more than one Nash equilibrium, couldn't we like, alternate? >> We could, and in fact you'll notice that in so far, I have not talked about the case where you have multiple Nash equilibrium. Because in fact, that's its own problem. If I have one Nash equilibrium, then we know what's going to happen. If I have two Nash equilibria, which one do we choose? Now, what's important to know here, is that if we have two Nash Equilibria then that means that if you're in any one of them, you won't leave it. So it's not like you'll move from one Nash Equilibrium to another, sort of in the general case. Because the fact that it's an equilibrium means that if everyone else is fixed except one person, that person won't choose to move. But. I haven't said anything about how you would choose among multiple Nash equilibria. And that's beyond the scope of this class. >> Okay. >> But it is an active researcher and in fact some of my own students have done some work in thinking about. What it would mean to choose and actually the answer always boils down to, let's not worry about that. >> [LAUGH] All right. But I'm still kind of disturbed by this, right? So it seems like it's sort of saying, is if we knew the world was going to end tomorrow then we might as well be greedy bastards(雜種). And, in fact, if we know that the world is going to end at any particular time, which of course it will. We might as well be greedy bastards. >> That's right. So, be greedy bastards. >> Oh, no! >> Or at least, or at least be you know, Nash bastards. >> [LAUGH] We'll, the Nash is kind of greedy right, it's like I'm always taking an action that's always best for me. >> Yes, you are. >> No. >> Every single thing we've talked about today. Has always assumed that you're always doing what's best for you. The fact that it might also be best for someone else is neither here nor there, you're always doing what's best for you. >> Right, I get, I get that, but it's, it sounds like this argument is saying that we might as well be like that, like all the time and never really form, like never, never self-sacrifice even a little bit for greater advantage even to yourself. >> Well, that's not true. This, it's just all hidden in the utilities, once you've gone through all the what I'm going to do today, what I'm going to do tomorrow, you add it all up, whatever the right strategy is to take. That's the strategy that you take because it's already accounted for all the self-sacrifice you might do that then leads to later good deeds or later good outcome. So the wonderful thing about all of this, is that this matrix is everything you need to know. It captures the future, past and everything else, so everything you need to know is right here. And it's all been reduced to a bunch of pairs of numbers. But you did say something kind of interesting though, Michael, which I think is worth pointing out, which is that everything you say requires knowing when, at least in this sort of iterated version, requires knowing when the world is going to end. So an interesting question to ask would be, what would happen if I knew the world was going to end but I didn't know when. Would that change my behavior. >> It, it doesn't seem like that should make any difference because it's still going to end. If not today then tomorrow. Or if not tomorrow then the next day. >> True. But I bet yeah it does make a difference. >> I'm willing to go and think about that. >> Okay, why don't you go think about it and let me know. >> Alright. I'll, I'll tell you about it in the next lecture then. >> Okay, I like that. Excellent. Okay, cool. So, let's move on. >> [LAUGH] Okay, fair enough.

GAME THEORY: WHAT HAVE WE LEARNED?

- STRATEGIES: PURE VS MIXED
- ANDREW MOORE
- PRISONER'S DILEMMA
- NASH!
- MECHANISM DESIGN
- WE ARE ALL IN THE MATRIX
- MINIMAX
- HIDDEN / PERFECT
- (NON) ZERO SUM
- (NON) DETERMINISTIC

-1, -1	-9, 0
0, -9	-6, -6

31. All right, Michael. So, I think that brings us to the end of what I wanted to talk about anyway. So, can you help me remember what it is that we've learned today? In particular, what you've learned today? >> Sure. So, I guess the first think I learned is that game theory can make you depressed. And, in fact, in particular that my friend Charles, given the opportunity. Would totally drop a dime on me just to save a month of incarceration. >> Yeah. >> Wait, no, no, no, I don't think you summarized that correctly. Game theory is not depressed. It's depressing. >> Oh, yeah. That's a good point, that's a good point. >> And Michael is not cruel. He is the victim of cruelty. >> I don't think so. Because you want to know what the secret here is, Michael? You've got a little matrix of numbers. Those number capture what's going on. The truth, Michael, is that, if we were in Prisoner's Dilemma. I would cooperate with you because my utility is not simply the number of months that I would spend in jail. But it's the number of months you also would spend in jail. >> Oo. Interesting. >> So if I, the best way to beat Prisoner's Dilemma, is to change the numbers. >> [LAUGH] I see. It's like the Kobayashi Maru of game theory. >> Exactly. So there's an interesting question for you right there, Michael. If I had prisoner's dilemma here, here, let's write it out so you can remember. If I had prisoner's dilemma here, we already know that we're going to end up here, because that's what the numbers tell us to do. But what we'd have to do is change the game. So how would you change the game in prisoner's dilemma? I see. So, if, if we're thinking about it in particular in terms of I care about how long you spend in jail. Maybe not as much as I care about how much I spend in jail. Like maybe half as much. >> Mm-hm. >> Then, the payments shift, right? >> Right. >> So, now we have like minus1 and a half, minus 1 and a half in the upper left hand corner. >> Mm-hm. Minus 9 comma minus 4.5, minus 4.5, minus 9 and minus 9, minus 9. >> So, yeah. >> So now that, that bottom right becomes a lot less attractive if we actually care about the other person. >> Right. >> Well that's, that's, that's, okay, I'm less depressed now. >> Except of course, that requires that you feel that way internally and that I feel that way internally. There's another way that you could change the game here. Which is, [what happens to sniches\(告密\) in jail?](#) >> They are rewarded. >> No. No they're not. >> [They're punished.](#) >> Yes. >> Oh. >> So, [if you a part of the criminal fraternity, and you don't like prisoners dilemma, then what](#)

you have to do is to create a system where the people who snitch get punished. So it's not just the months that they spend in jail, it's everything else that's going to happen to them if they drop a dime. >> So you're saying that minus 6, minus 6, ends up being worse? >> No, what I'm. >> No, the minus, wait, no, wait, what? Yeah. Oh, the zero ends up getting, oh I see the zero ends up being worse. Because even though your not in jail your going to get I don't know somehow thwarted or, or punished for your past behaviors. >> Accosted. >> Interesting. >> That's right. So that's what you have to do and that works not just with criminals but with the real world. Whenever your in this sort of situation like a prisoners dilemma, you can change the game by changing everyone's utilities. Like for example hiring police officer, police officers or hiring members of the mob to take care of everything. >> I see. So it almost seems like what you're talking about is a kind of inverse game theory, where if there's a particular behavior that I want to see. How do I set up the payments and rewards so that that behavior is encouraged. >> Right, and by the way that has a name, and it's called mechanism design. Mechanism design? >> Yes. >> I'm not sure I understand either of those words. [LAUGH] >> Well, that's where you're trying to set up the set of incentives, the mechanisms that you're using to pay people. You're trying to design them in such a way to get particular behavior. This is what a lot of economics is all about. This is what a lot of government is all about. Tax breaks for example, for mortgage interest, encourages you to buy a home, rather than rent a home. >> I see, by changing the payoff structure. >> Right. >> Oh that's neat. >> And so that's what we learned today. At least right now. [LAUGH] >> Okay. Alright. So, let's see. So, just to try to rattle off some of the other things. The whole notion of Game Theory. We talked about, especially the idea that. You can think about a game as a tree or you could represent it as a matrix. And, I believe you said, repeatedly, the matrix has everything. Is that how you said it? Or the matrix is all you need. >> Yip! Let's see. We talked about minimax and maximin. >> Mm-hm. >> We, we relaxed a bunch of constraint on games. >> Mm-hm. >> So we, we looked at both perfect and hidden information. >> Mm-hm. >> We looked at both zero sum and non zero sum. >> We learned a lot today. We looked at deterministic and [UNKNOWN] I would want to say, but you called it non-deterministic. And assuming that we can get rid of the first two bullet items that look like maybe they were jokes. I would suggest saying things like, we talked about what strategies are and that they come in different flavors. We talked about the evil prisoners dilemma game. >> Mm-hm. What else? You gotta give me more, otherwise it's [INAUDIBLE]. >> Oh, more, good point. Andrew Moore gave lots of really good examples that >> yes. >> Michael may be cruel, but Andrew Moore is awesome. >> He's more cool than me. Andrew Moor is very cool. All of the examples that we've used today, or almost all of the examples we used today actually come from Andrew Moore's slides. Andrew Moore is a professor at Carnegie Mellon or at least he was before he went off to Google. And is a really smart guy who cares very much about machine learning. And game theory and produced a bunch of slides that it turns out lots and lots of people use in their own courses. And his examples were so good for game theory. That I decided to coop them with his permission of course. He tells everyone that they may use them. And, in fact, we have pointers to the slides in the resources links and folders for all of you to look at. And I recommend that you do. Did we learn anything else, Michael? >> The only other thing that I would want to mention is NASH which is a concept that is nashtastic. >> It is nashtastic. There are by the way, I should mention briefly. Other kinds of equilibria concepts they're beyond the scope of this class. But there's a whole lot more to game theory as you might imagine. And sometimes when they ask you to in these situations where you can't do what you want to do, you end up in these prisoner's lemonade situations. Other kinds of equilibria can get you out of it. And I'm going to argue without explaining why that the way that they get around this is by introducing Other ways of doing various kinds of communication. and, in fact, I claim they're a particular part of mechanism design. But that's a topic for another day. >> Okay. Fair enough. >> Okay. Did we learn anything else, Michael? >> I don't know. That's what I was thinking about. I mean, that seems like a lot to absorb. And, the other thing is that repeated games, even the prisoner's dilemma kind of unravel if you know when the end. And I was going to look into what happens if you don't

know when they're going to end. >> Okay. So, I guess, that will be something that we will learn. Next time. >> Right, what we, what will we have will learned? >> Yes. Future past tense. >> [LAUGH] >> Alright, Michael, well, I think that's about it. At least my brain is full. So I will talk to you next time, and you get to lead what I believe is the last full lesson of the course. >> Oh, exciting. It is exciting. Alright, well bye Michael. You have fun. I'll see you next time. >> Alright. Bye Charles. >> Bye.