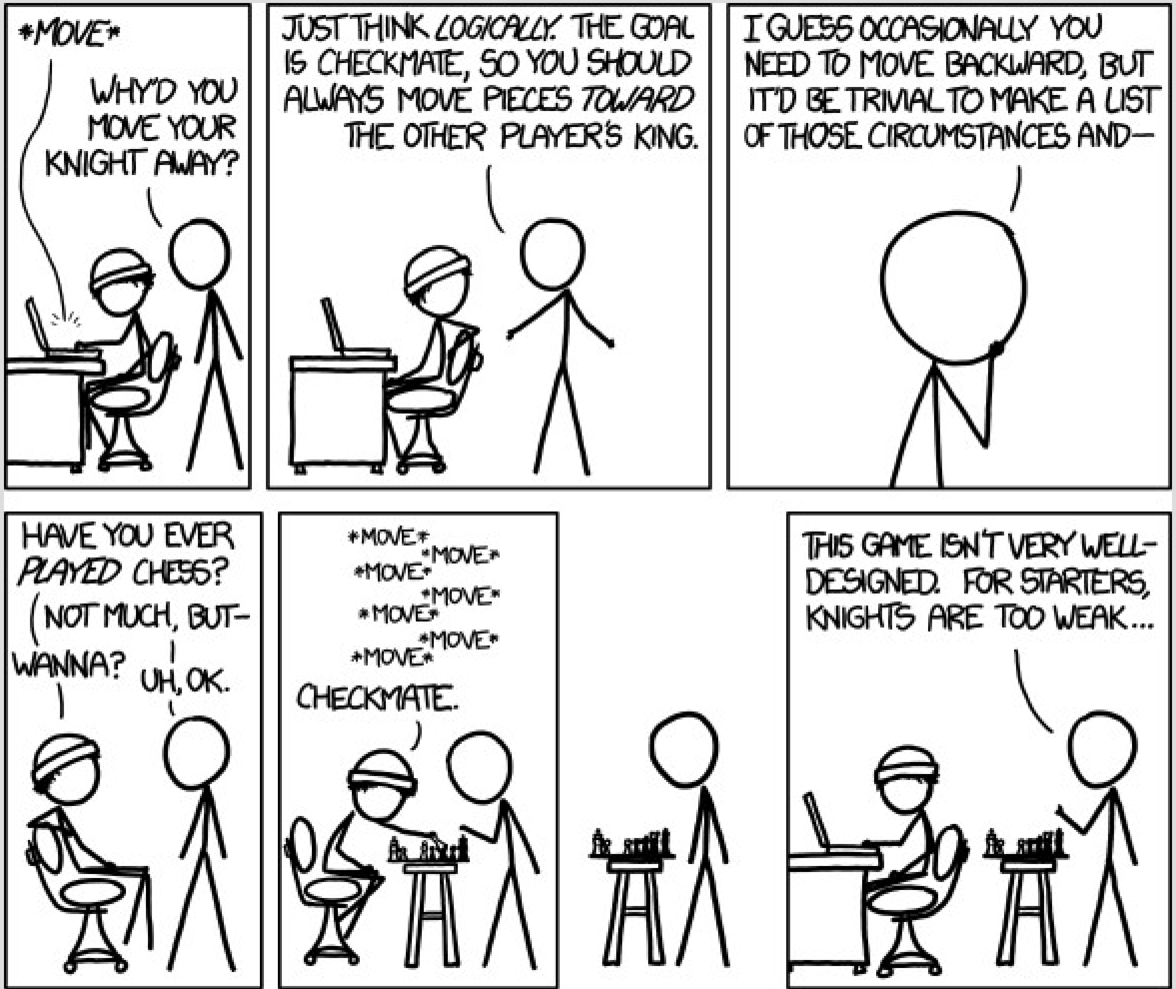
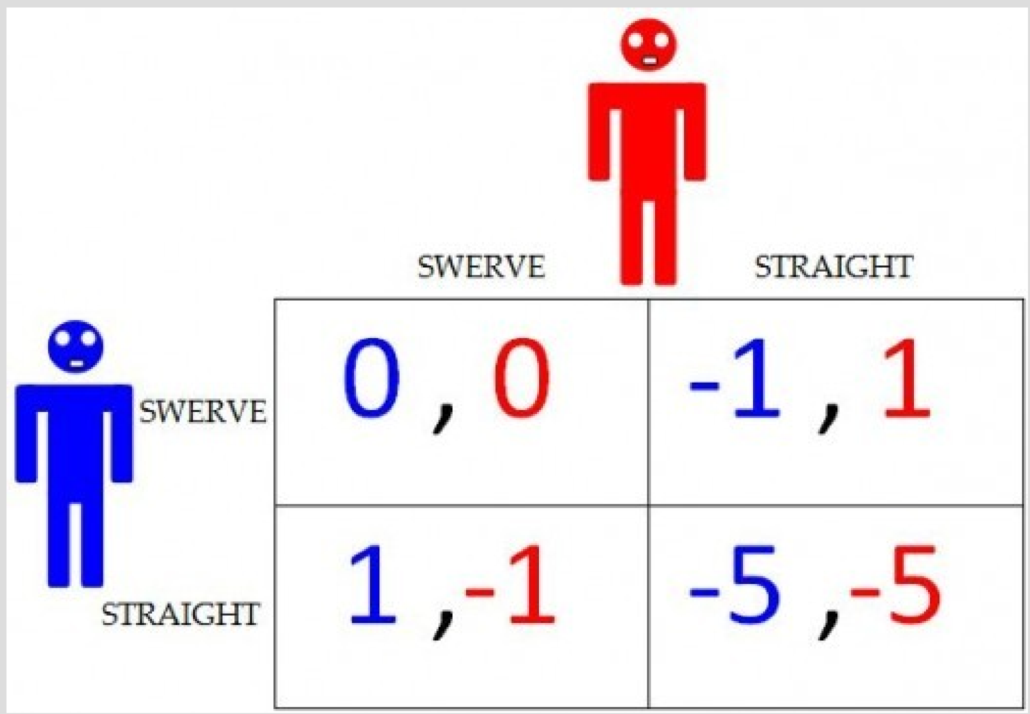
Game Theory (Ch 17.5)



Game theory

Typically game theory uses a payoff matrix to represent the value of actions



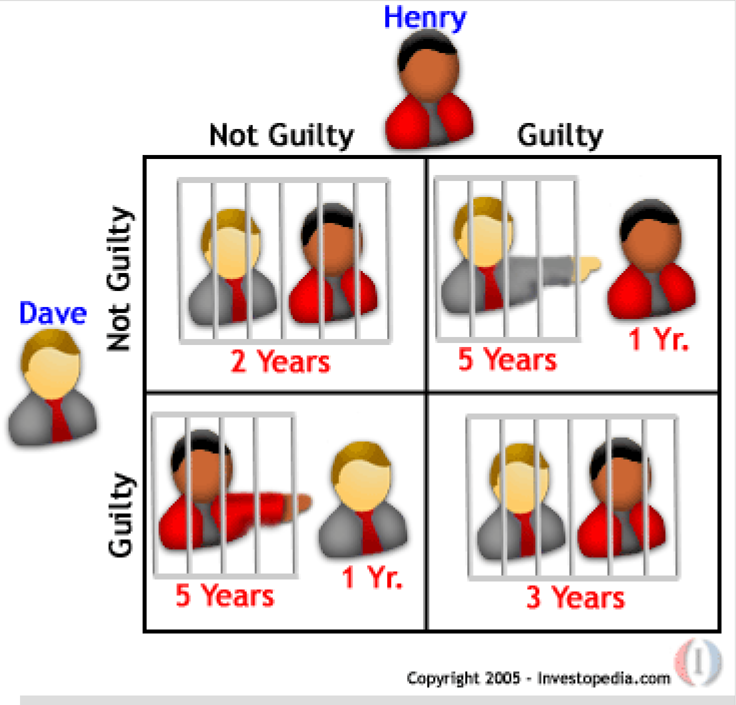
The first value is the reward for the left player, right for top (positive is good for both)

Dominance & equilibrium

Here is the famous “prisoner's dilemma”

Each player chooses one action without knowing the other's and the is only played once

PRISONER 1 \ PRISONER 2	Confess	Lie
Confess	-8 , -8	0 , -10
Lie	-10 , 0	-1 , -1

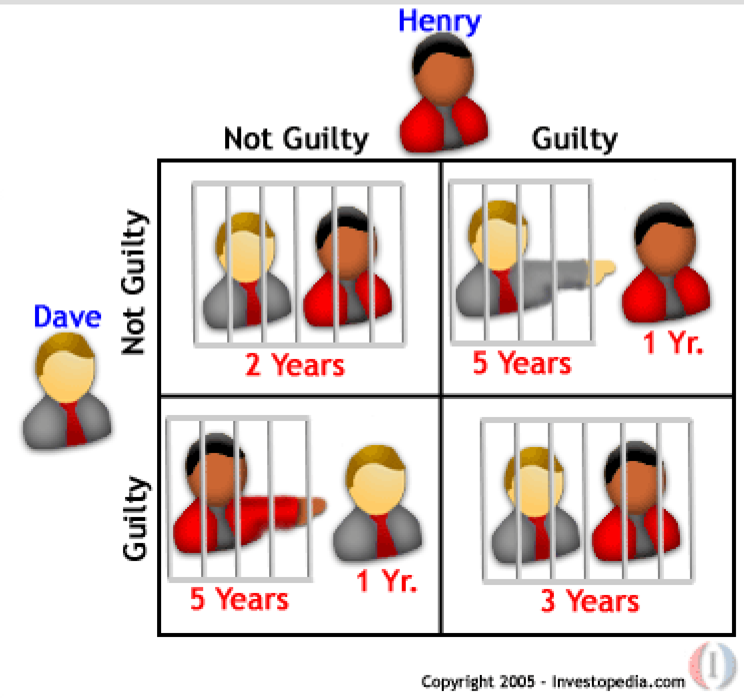


Dominance & equilibrium

What option would you pick?

Why?

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1



Dominance & equilibrium

What would a rational agent pick?

If prisoner 2 confesses, we are in the first column... -8 if we confess, or -10 if we lie
--> Thus we should confess

If prisoner 2 lies, we are in the second column,
0 if we confess,
-1 if we lie
--> We should confess

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1

Dominance & equilibrium

It turns out regardless of the other player's action, it is in our personal interest to confess

This is the Nash equilibrium, as any deviation of strategy (i.e. lying) can result in a lower score (i.e. if opponent confesses)

The Nash equilibrium looks at the worst case and is greedy

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1

Dominance & equilibrium

Formally, a Nash equilibrium is when the combined strategies of all players give no incentive for any single player to change

In other words, if any single person decides to change strategies, they cannot improve

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1

Dominance & equilibrium

Alternatively, a <u>Pareto optimum</u> is a state where no other state can result in a gain or tie for all players (excluding all ties)

If the PD game, [-8, -8] is a Nash equilibrium, but is not a Pareto optimum (as [-1, -1] better for both players)

However [-10,0] is also a Pareto optimum...

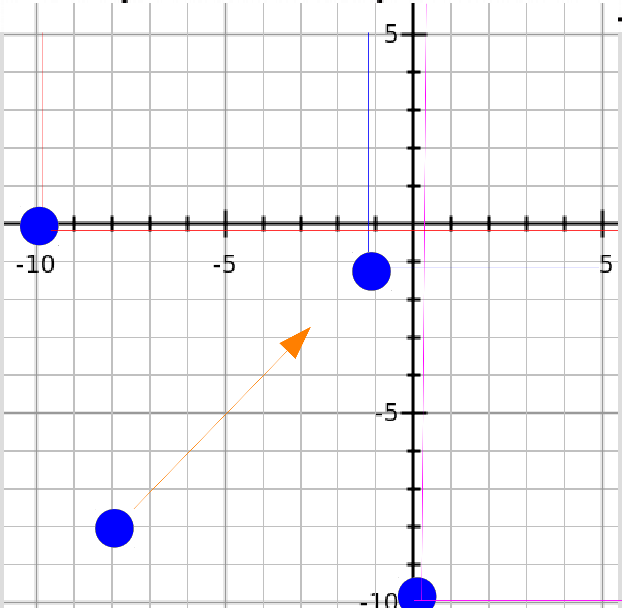
		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1

Dominance & equilibrium

To find <u>Pareto optimum</u>, you can simply graph all the points (x-axis = p1, y-axis = p2)

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8 , -8	0 , -10
	Lie	-10 , 0	-1 , -1

Any points that no other points up and to the right are <u>Pareto optimum</u>



The only point with something up&right is (-8,-8) (orange line shows (-1,-1) better for both)

Dominance & equilibrium

Every game has at least one Nash equilibrium and Pareto optimum, however...

- Nash equilibrium might not be the best outcome for all players (like PD game, assumes no cooperation)

- A Pareto optimum might not be stable (in PD the [-10,0] is unstable as player 1 wants to switch off “lie” and to “confess” if they play again or know strategy)

Dominance & equilibrium

Find the Nash and Pareto for the following:
(about lecturing in a certain csci class)

Student

Teacher	pay attention	sleep
prepare well	5, 5	-2, 2
slack off	1, -5	0, 0

Find best strategy

Another way to find a Nash equilibrium?

If it is zero-sum game, can use minimax
as neither player wants to switch for Nash
(our PD example was not zero sum)

Let's play a simple number game: two players
write down either 1 or 0 then show each other.
If the sum is odd, player one wins. Otherwise,
player 2 wins (on even sum)

Find best strategy

This gives the following payoffs:

Player 1 \ Player 2	Pick 0	Pick 1
Pick 0	-1, 1	1, -1
Pick 1	1, -1	-1, 1

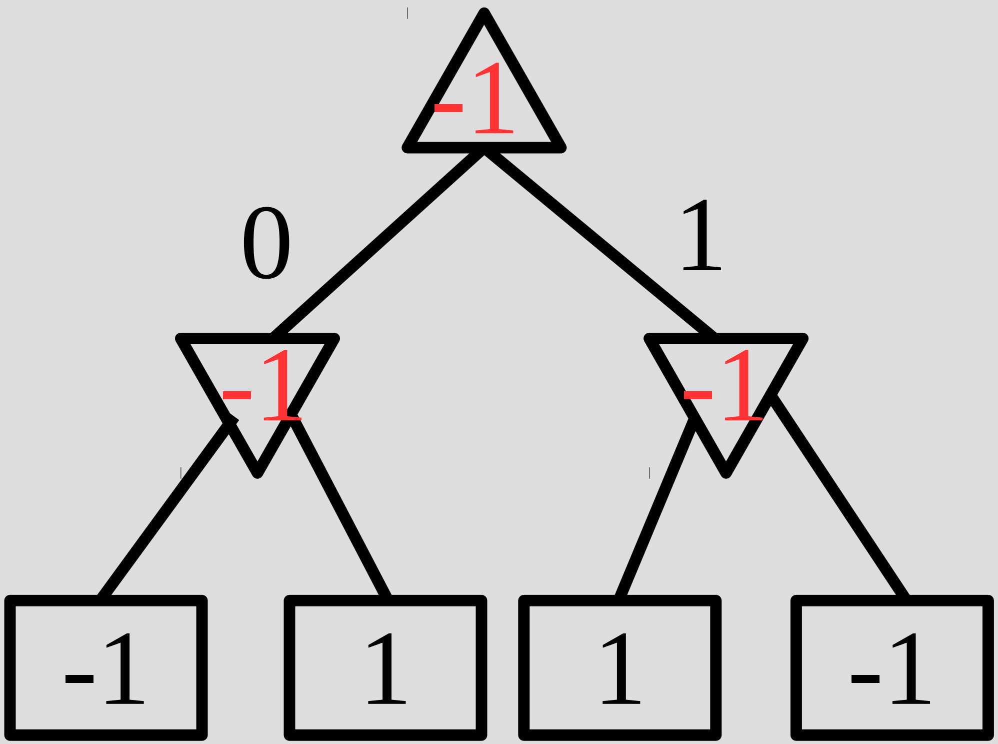
(player 1's value first, then player 2's value)

We will run minimax on this tree twice:

1. Once with player 1 knowing player 2's move (i.e. choosing after them)
2. Once with player 2 knowing player 1's move

Find best strategy

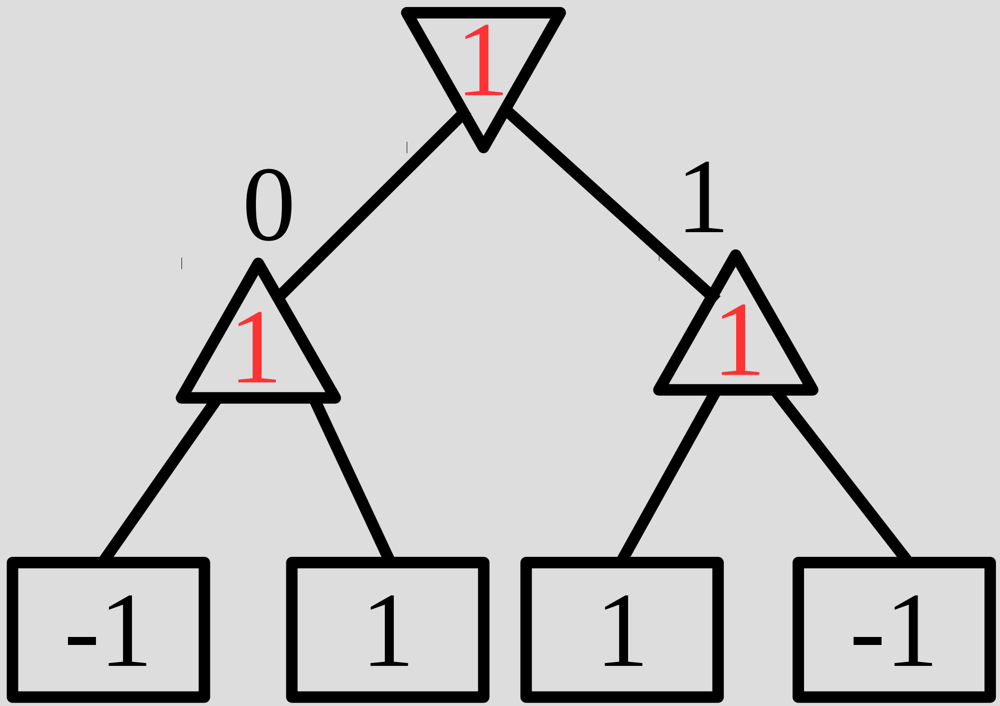
Player 1 to go first (max):



If player 1 goes first, it will always lose

Find best strategy

Player 2 to go first (min):



If player 2 goes first, it will always lose

Find best strategy

This is not useful, and only really tells us that the best strategy is between -1 and 1 (which is fairly obvious)

This minimax strategy can only find pure strategies (i.e. you should play a single move 100% of the time)

To find a “mixed strategy” (probabilistically play), we need to turn to linear programming

Find best strategy

A <u>pure strategy</u> is one where a player always picks the same strategy (deterministic)

A <u>mixed strategy</u> is when a player chooses actions probabilistically from a fixed probability distribution (i.e. the percent of time they pick an action is fixed)

If one strategy is better or equal to all others across all responses, it is a <u>dominant strategy</u>

Find best strategy

The definition of a Nash equilibrium is when no one has an incentive to change the combined strategy between all players

So we will only consider our opponent's rewards (and not consider our own)

This is a bit weird since we are not considering our own rewards at all, which is why the Nash equilibrium is sometimes criticized

Find best strategy

First we parameterize this and make the tree stochastic:

Player 1 will choose action “0” with probability p, and action “1” with (1-p)

If player 2 always picks 0, so the payoff for p2:
$(1)p + (-1)(1-p)$
If player 2 always picks 1, so the payoff for p2:
$(-1)p + (1)(1-p)$

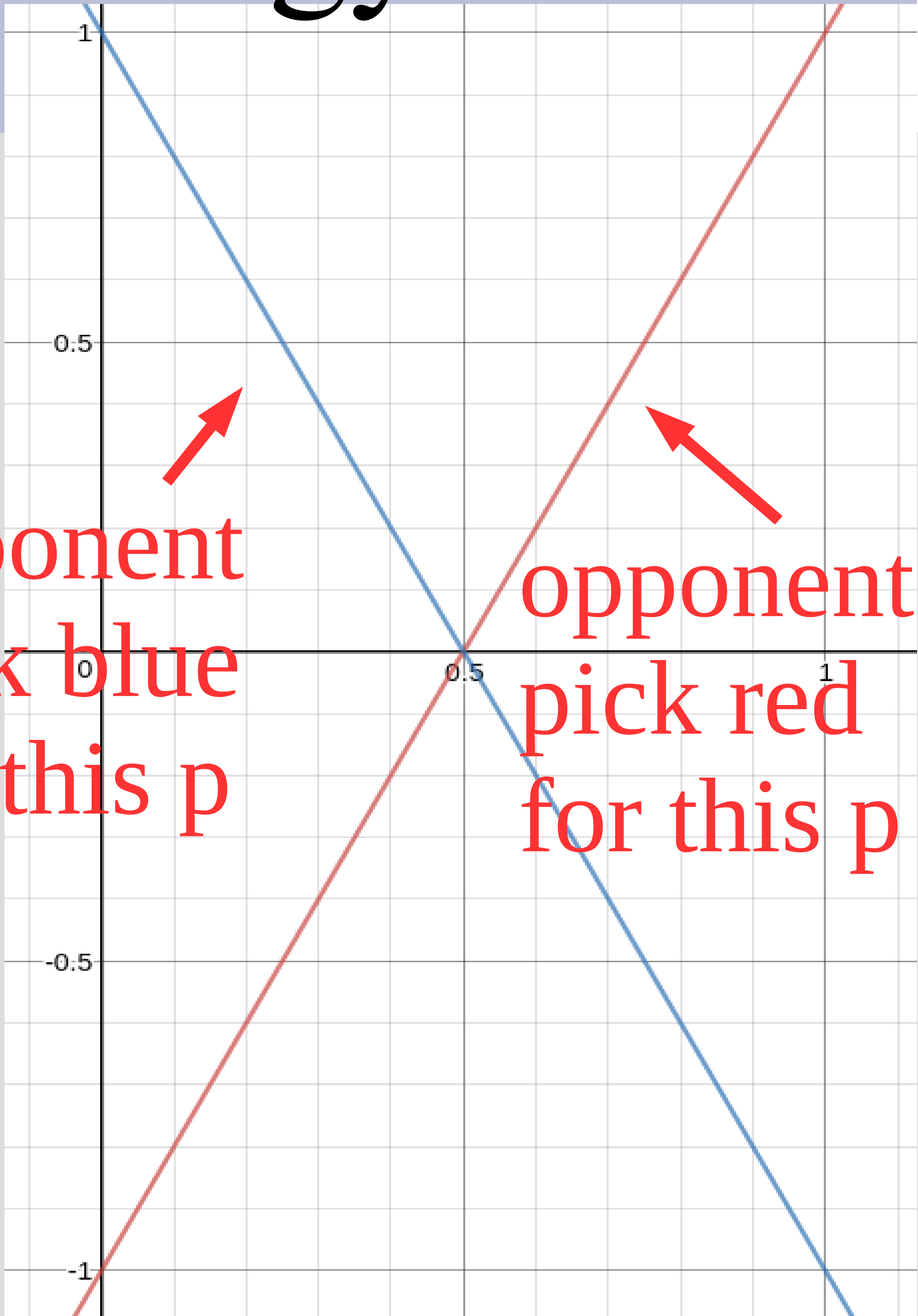
Find best strategy

Plot these two lines:

$$U = (1)p + (-1)(1-p)$$
$$U = (-1)p + (1)(1-p)$$

As we maximize, the opponent gets to pick which line to play

Thus we choose the intersection



Find best strategy

Thus we find that our best strategy is to play 0 half the time and 1 the other half

The result is we win as much as we lose on average, and the overall game result is 0

Player 2 can find their strategy in this method as well, and will get the same 50/50 strategy (this is not always the case that both players play the same for Nash)

Find best strategy

How does this compare on PD?

	Confess	Lie
Confess	-8 , -8	0 , -10
Lie	-10 , 0	-1 , -1

Player 1: p = prob confess...

P2 Confesses: -8*p + 0*(1-p)

P2 Lies: -10*p + (-1)*(1-p)

Cross at negative p, but red line is better (confess)