CS 440/ECE448 Lecture 31: Mixed Equilibrium

Mark Hasegawa-Johnson



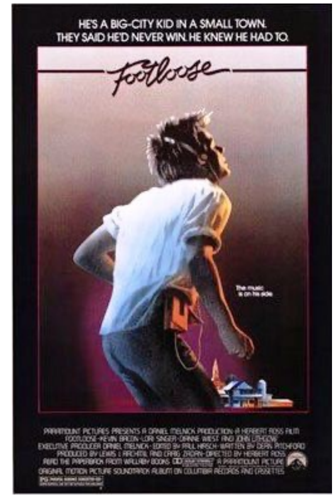

Images:
Jyotsnu Vasudeva
https://commons.wikimedia.org/wiki/File:Megaceryle_alcyon_f_Sam_Smith_Park_Toronto.jpg
https://commons.wikimedia.org/wiki/File:Paparazzi_La_Dolce_Vita.jpg
Alejandro Aragon

Outline

- The Game of Chicken
- Mixed equilibrium: Randomness is rational behavior
- Adversarial learning
- Generative adversarial networks

Game of Chicken

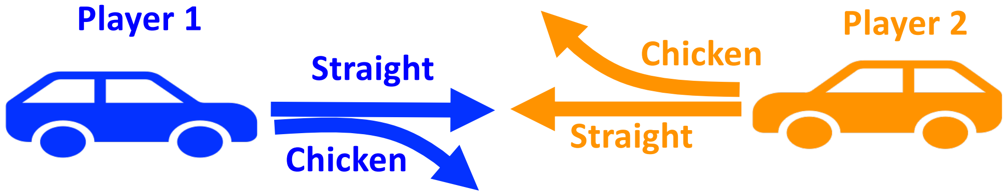
Filmmakers and storytellers love the game of chicken. Why?



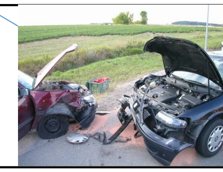
- Two players each claim to be braver than the other
- To prove their bravery, they each claim willingness to do something stupid (drive toward a cliff, drive directly toward one another, try to sabotage one another's relationship with Henry Golding…)
- Outcomes:
  - If one player chickens out, the chicken loses, the other wins
  - If both players chicken out, nobody wins or loses
  - If neither player chickens out, they both LOSE BIG

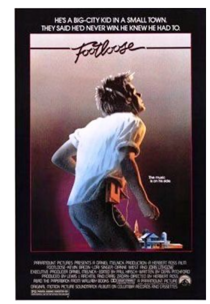
http://en.wikipedia.org/wiki/Game_of_chicken

Game of Chicken



- Example payoff matrix:
  - If one player chickens out, he loses $1000, and the other wins $2000
  - If both players chicken out, they each win $1000
  - If neither player chickens out, they both lose $10,000 (the cost of the car)

Bob \ Alice	Defect	Cooperate
Defect		Bob: 2, Alice: -1
Cooperate	Bob: -1, Alice: 2	Bob: 1, Alice: 1



http://en.wikipedia.org/wiki/Game_of_chicken

Prisoner’s Dilemma vs. Game of Chicken

Prisoner’s Dilemma

	Defect	**Cooperate**
Defect	**Lose** / **Lose**	**Lose Big** / **Win Big**
Cooperate	**Win Big** / **Lose Big**	**Win** / **Win**

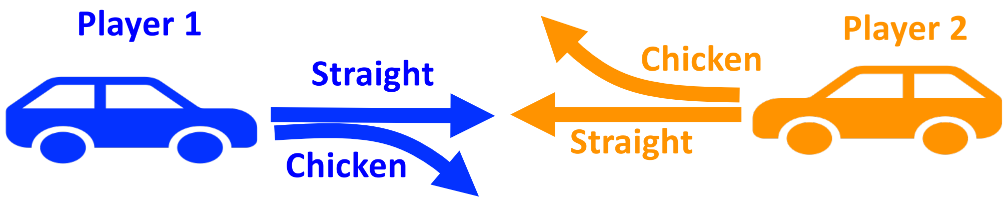
Players cut their losses by defecting if the other player defects

Game of Chicken

	Defect	**Cooperate**
Defect	**Lose Big** / **Lose Big**	**Lose** / **Win Big**
Cooperate	**Win Big** / **Lose**	**Win** / **Win**

Defecting, if the other player defects, is the worst thing you can do

Game of Chicken

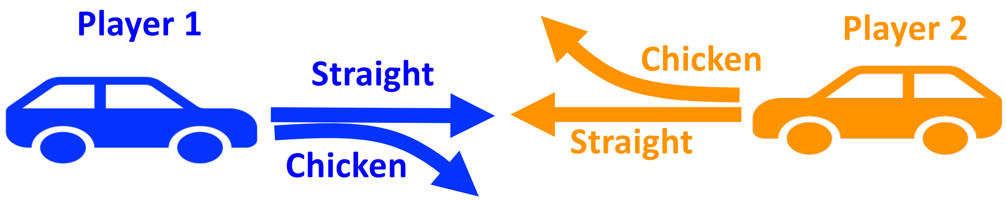


		Alice	
		Defect	Cooperate
Bob	Defect	-10 / -10	2 / -1
	Cooperate	-1 / 2	1 / 1

- No dominant strategies for either player
  - If the other player defects, I should cooperate!
  - If the other player cooperates, I should defect!

http://en.wikipedia.org/wiki/Game_of_chicken

Game of Chicken

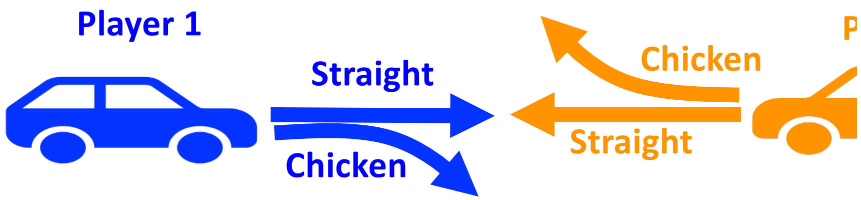


Bob \ Alice	Defect	Cooperate
Defect	-10 / -10	-1 / 2
Cooperate	2 / -1	1 / 1

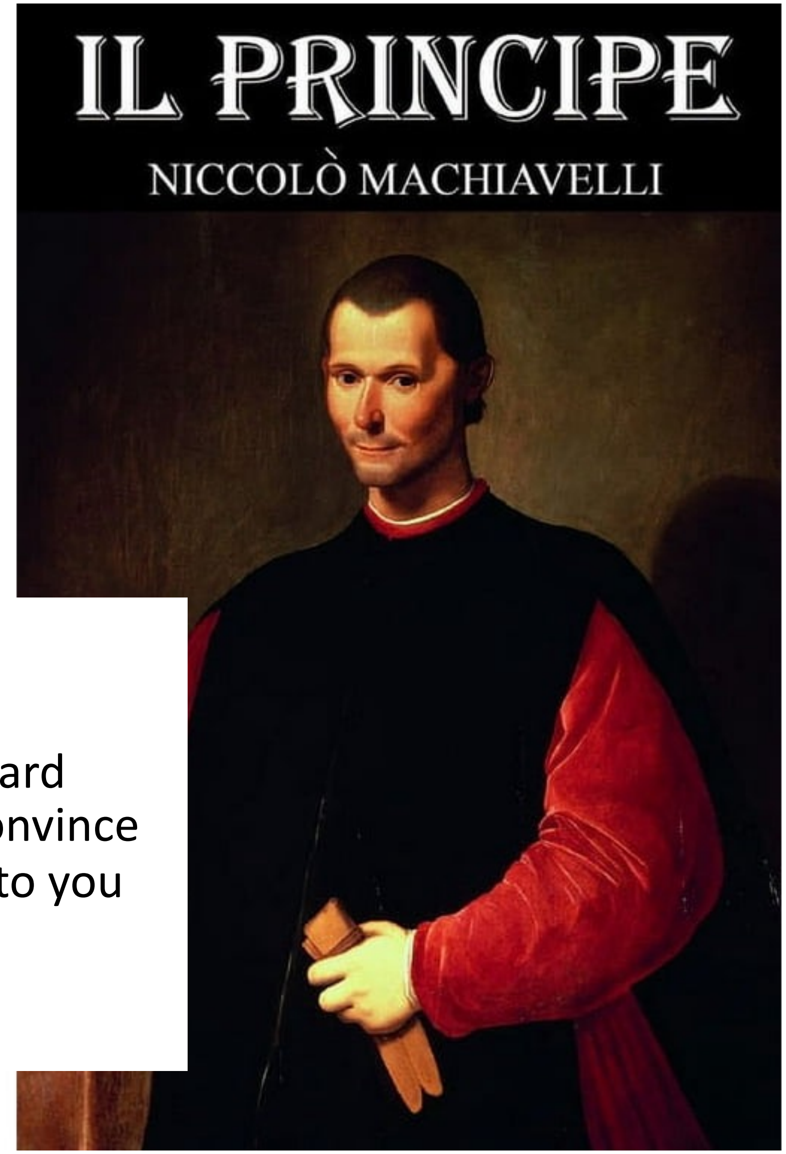
- There are two Nash equilibria: (C,D) and (D,C)
  - The winner is not motivated to change
  - The loser can’t win by changing unilaterally

http://en.wikipedia.org/wiki/Game_of_chicken

Game of Chicken



- How can you win?
  - Bluff!
  - Convince the other player that your personal reward function is different from the one shown here: Convince them that the physical loss would be less painful to you than allowing the other player to win
  - In words: Pretend to be irrational



http://en.wikipedia.org/w

Outline

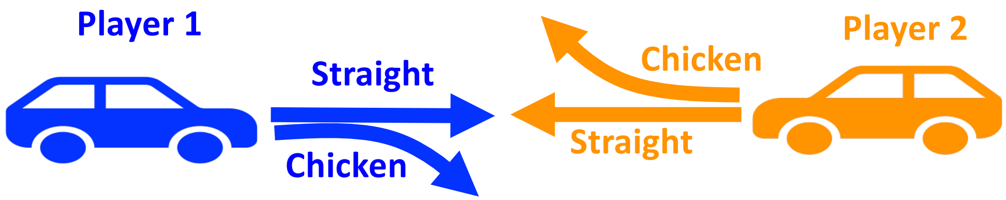
- The Game of Chicken
- Mixed equilibrium: Randomness is rational behavior
- Adversarial learning
- Generative adversarial networks

Irrational versus Random

The game of chicken has two different types of Nash equilibria:

- **Be irrational**: Bluff. One player convinces the other that he or she will behave irrationally. The other player concedes the game. Result: (straight,chicken) or (chicken,straight).
- **Be random:** Mixed Nash Equilibrium.
  - Alice chooses a move at random, according to some probability distribution. She tells Bob, in advance, what probability distribution she will use.
  - Bob responds rationally.
  - One of Bob's rational options is to choose his move, also, at random.

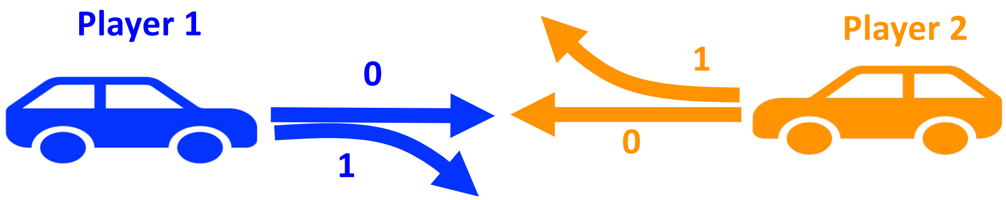
Mixed strategy = Random



		Alice: Straight	Alice: Chicken
Bob	Straight	Bob: -10, Alice: -10	Bob: 2, Alice: -1
	Chicken	Bob: -1, Alice: 2	Bob: 1, Alice: 1

- **Mixed strategy:** a player chooses between the different possible actions according to a probability distribution.
- **Mixed Nash equilibrium**:
  - One or both players choose their actions ***at random***
  - For both players, this is a ***rational*** thing to do, i.e., they can’t increase their expected reward by using a non-random strategy

Mixed strategy = Random



Bob \ Alice	0	1
0	Alice: -10, Bob: -10	Alice: -1, Bob: 2
1	Alice: 2, Bob: -1	Alice: 1, Bob: 1

For example, suppose that both players, independently, decide to defect (go straight, i.e., action 0) with probability 1/10.

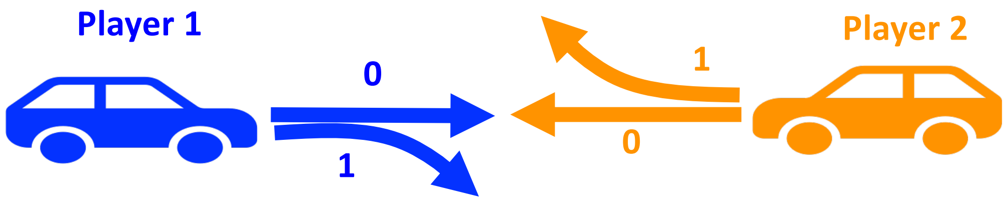
- Random variable A: Alice's action

$$P(A = 0) = \frac{1}{10}, \qquad P(A = 1) = \frac{9}{10}$$

- Random variable B: Bob's action

$$P(B = 0) = \frac{1}{10}, \qquad P(B = 1) = \frac{9}{10}$$

Notation for mixed strategies



Alice \ Bob	0	1
0	-10 (Alice), -10 (Bob)	2 (Alice), -1 (Bob)
1	-1 (Alice), 2 (Bob)	1 (Alice), 1 (Bob)

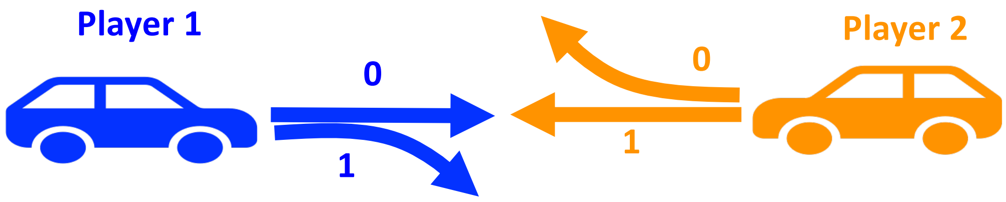
- Let's say that each player chooses an action according to these probabilities:

$$\boldsymbol{\alpha} = \begin{bmatrix}\alpha_0\\ \alpha_1\end{bmatrix} = \begin{bmatrix}P(A=0)\\ P(A=1)\end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \begin{bmatrix}P(B=0)\\ P(B=1)\end{bmatrix}$$

- …and gets these rewards as a result:

$$\boldsymbol{A} = \begin{bmatrix}A_{0,0} & A_{0,1}\\ A_{1,0} & A_{1,1}\end{bmatrix} = \begin{bmatrix}-10 & 2\\ -1 & 1\end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix}B_{0,0} & B_{0,1}\\ B_{1,0} & B_{1,1}\end{bmatrix} = \begin{bmatrix}-10 & -1\\ 2 & 1\end{bmatrix}$$

Notation for mixed strategies



Alice \ Bob	0	1
0	Alice: -10, Bob: -10	Alice: 2, Bob: -1
1	Alice: -1, Bob: 2	Alice: 1, Bob: 1

Using this notation, we can calculate the expected rewards for Alice as a function of her action:

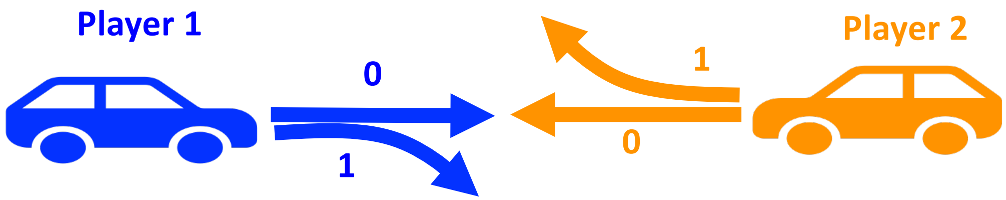
$$E[r_A|\mathrm{A}=0] = A_{0,0}\beta_0 + A_{0,1}\beta_1 = \left(\frac{1}{10}\right)(-10) + \left(\frac{9}{10}\right)(2) = \frac{8}{10}$$

$$E[r_A|\mathrm{A}=1] = A_{1,0}\beta_0 + A_{1,1}\beta_1 = \left(\frac{1}{10}\right)(-1) + \left(\frac{9}{10}\right)(1) = \frac{8}{10}$$

Putting that in matrix form, we get that

$$E[\boldsymbol{r}_A] = \begin{bmatrix} E[r_A|\mathrm{A}=1] \\ E[r_A|\mathrm{A}=0] \end{bmatrix} = \boldsymbol{A\beta} = \begin{bmatrix} -10 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{10} \\ \frac{9}{10} \end{bmatrix} = \begin{bmatrix} \frac{8}{10} \\ \frac{8}{10} \end{bmatrix}$$

Notation for mixed strategies



Alice \ Bob	0	1
0	-10 (Alice), -10 (Bob)	2 (Alice), -1 (Bob)
1	-1 (Alice), 2 (Bob)	1 (Alice), 1 (Bob)

Similarly, the expected rewards for Bob as a function of his action are:

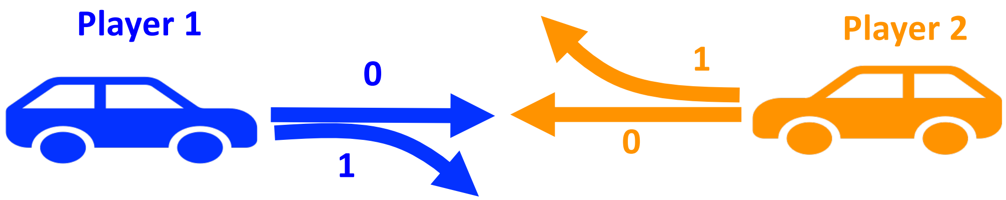
$$E[r_B|\mathrm{B}=0] = \alpha_0 B_{0,0} + \alpha_1 B_{1,0} = \left(\frac{1}{10}\right)(-10) + \left(\frac{9}{10}\right)(2) = \frac{8}{10}$$

$$E[r_B|\mathrm{B}=1] = \alpha_0 B_{1,0} + \alpha_1 B_{1,1} = \left(\frac{1}{10}\right)(-1) + \left(\frac{9}{10}\right)(1) = \frac{8}{10}$$

Putting that in matrix form, we get that

$$E[\boldsymbol{r}_B] = [E[r_B|\mathrm{B}=0], E[r_B|\mathrm{B}=1]] = \boldsymbol{\alpha}^T\boldsymbol{B} = \left[\frac{1}{10}, \frac{9}{10}\right]\begin{bmatrix}-10 & -1\\ 2 & 1\end{bmatrix} = \left[\frac{8}{10}, \frac{8}{10}\right]$$

Mixed Nash equilibrium



Bob \ Alice	0	1
0	Alice: -10, Bob: -10	Alice: -1, Bob: 2
1	Alice: 2, Bob: -1	Alice: 1, Bob: 1

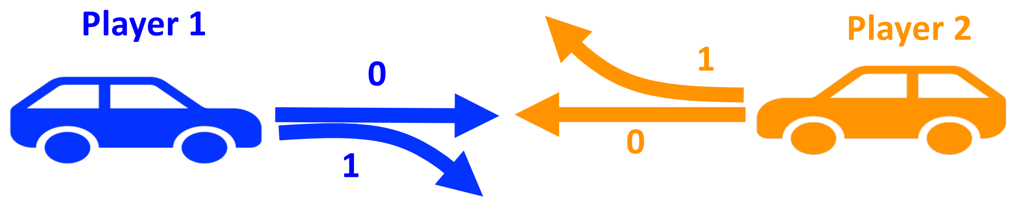
- Neither player can increase their reward by choosing a deterministic action:

$$E[r_A|\mathrm{A}=0]=E[r_A|\mathrm{A}=1]=\frac{8}{10}$$

$$E[r_B|\mathrm{B}=0]=E[r_B|\mathrm{B}=1]=\frac{8}{10}$$

- Going straight, chickening out, or choosing at random between the two are all rational actions. Random choice with probability $\mathrm{P}(1)=\frac{9}{10}$ is a rational action, as would be any other strategy.

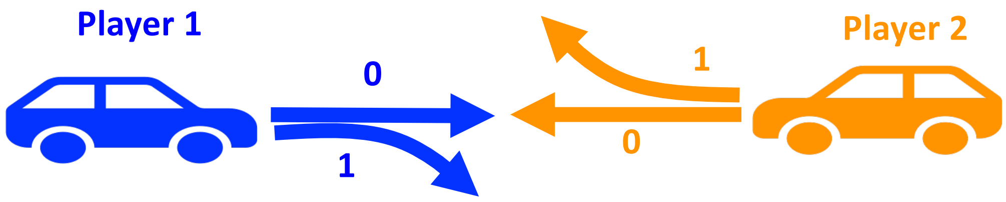
Unstable equilibrium



		Alice 0	Alice 1
Bob	0	-10 (Alice), -10 (Bob)	-1 (Alice), 2 (Bob)
Bob	1	2 (Alice), -1 (Bob)	1 (Alice), 1 (Bob)

- Random choice with probability $\mathrm{P}(A=1)=\frac{9}{10}$ is a rational action; but any other strategy would also be rational. So why should Alice choose this strategy in particular?
- Notice: if Alice increases her probability of going straight ***even a little bit***, then it suddenly becomes rational for Bob to always chicken out.
- The mixed Nash equilibrium in this case is an ***unstable equilibrium***: as long as $\mathrm{P}(A=1)=\mathrm{P}(B=1)=\frac{9}{10}$, then there is no reason for either of them to change, but if one of them changes by even a little bit, then the other is forced to change in response (if behaving rationally).

Calculating the mixed equilibrium



Bob \ Alice	0	1
0	Alice: -10, Bob: -10	Alice: -1, Bob: 2
1	Alice: 2, Bob: -1	Alice: 1, Bob: 1

There is a Mixed Nash equilibrium if and only if there are some probabilities $0 \leq \alpha_1 \leq 1$ and $0 \leq \beta_1 \leq 1$ such that Alice's expected reward is independent of her action:

$$E[r_A|\mathrm{A}=0] = E[r_A|\mathrm{A}=1], \qquad \text{i.e.,} \qquad \begin{bmatrix} -1 \\ 1 \end{bmatrix}^T \boldsymbol{A}\boldsymbol{\beta} = \mathbf{0}$$

...and Bob's expected reward is independent of his action:

$$E[r_B|\mathrm{B}=0] = E[r_B|\mathrm{B}=1], \qquad \text{i.e.,} \qquad \boldsymbol{\alpha}^T \boldsymbol{B} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0$$

Does every game have a mixed-strategy equilibrium?

There is a Mixed Nash equilibrium iff $\exists \alpha_1 \in [0,1], \beta_1 \in [0,1]$ such that $E[r_A|\mathrm{A}=0] = E[r_A|\mathrm{A}=1]$ and $E[r_B|\mathrm{B}=0] = E[r_B|\mathrm{B}=1]$, i.e.,

$$\begin{bmatrix}-1\\1\end{bmatrix}^T \boldsymbol{A\beta} = \mathbf{0}$$

$$\boldsymbol{\alpha}^T \boldsymbol{B} \begin{bmatrix}-1\\1\end{bmatrix} = 0$$

- That's not true for every game.
- Some games have ONLY pure-strategy mixed equilibria.
- Some games have ONLY mixed-strategy equilibria.
- Some have both.
- … but every game at least one equilibrium (mixed, or pure, or both).

Does every game have a mixed-strategy equilibrium?

There is a Mixed Nash equilibrium iff $\exists\alpha_1 \in [0,1], \beta_1 \in [0,1]$ such that $E[r_A|\text{A}=0] = E[r_A|\text{A}=1]$ and $E[r_B|\text{B}=0] = E[r_B|\text{B}=1]$, i.e.,

$$\begin{bmatrix}-1\\1\end{bmatrix}^T \boldsymbol{A\beta} = \boldsymbol{0}$$

$$\boldsymbol{\alpha}^T \boldsymbol{B}\begin{bmatrix}-1\\1\end{bmatrix} = 0$$

- Prisoner's Dilemma has only one fixed-strategy Nash equilibrium (both players defect).
- Stag Hunt has:
  - 2 fixed-strategy Nash equilibria (either both players cooperate, or both players defect)
  - 1 mixed-strategy equilibrium (each player cooperates with probability 1/10).
- The Game of Chicken has:
  - 2 fixed strategy Nash equilibria (Alice defects while Bob cooperates, or vice versa)
  - 1 mixed-strategy Nash equilibrium (each player cooperates with probability 9/10).
- The Paparazzi game has only one mixed-strategy Nash equilibrium.

Try the quiz!

Try the quiz!

Outline

- The Game of Chicken
- Mixed equilibrium: Randomness is rational behavior
- Adversarial learning
- Generative adversarial networks

The Paparazzi game

- Alice is a famous movie star. Her agent announces that she will be at Illini Union signing autographs all day, but secretly, she might defect by going to Grainger to get some work done.
- Bob is a paparazzo. His job is to get Alice's photograph.
- If Alice and Bob are in the same location, Alice loses (-1), Bob wins (+1)
- If they are in different locations, Alice wins (+1), Bob loses (-1)

Alice \ Bob	0	1
0	Alice: -1, Bob: 1	Alice: 1, Bob: -1
1	Alice: 1, Bob: -1	Alice: -1, Bob: 1

The Paparazzi game

- Let's define Alice's cooperation probability to be the logistic sigmoid of some hidden scalar logit called a:

$$\boldsymbol{\alpha} = \begin{bmatrix} P(A=0) \\ P(A=1) \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} 1-\sigma(a) \\ \sigma(a) \end{bmatrix}$$

- Bob's cooperation probability is the sigmoid of a logit called b:

$$\boldsymbol{\beta} = \begin{bmatrix} P(B=0) \\ P(B=1) \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1-\sigma(b) \\ \sigma(b) \end{bmatrix}$$

The Nash Equilibrium

- Alice's expected rewards for defecting vs cooperating are
$$\begin{bmatrix} E[r_A|A=0] \\ E[r_A|A=1] \end{bmatrix} = \boldsymbol{A}\boldsymbol{\beta} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1-\sigma(b) \\ \sigma(b) \end{bmatrix} = \begin{bmatrix} 2\sigma(b)-1 \\ 1-2\sigma(b) \end{bmatrix}$$
- Alice gets equal reward from either action (and it is therefore rational for her to behave randomly) if $\sigma(b) = \frac{1}{2}$, i.e., if Bob is equally likely to either cooperate or defect.
- Bob's expected rewards are
$$\big[E[r_B|B=0], E[r_B|B=1]\big] = \boldsymbol{\alpha}^T \boldsymbol{B} = \begin{bmatrix} 1-\sigma(a) \\ \sigma(a) \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
$$= [1-2\sigma(a), 2\sigma(a)-1]$$
- Bob gets equal reward from either action (and it is therefore rational for him to behave randomly) if $\sigma(a) = \frac{1}{2}$, i.e., if Alice is equally likely to either cooperate or defect.

Could we use machine learning to find the Nash equilibrium?

- Suppose both Alice and Bob are using mixed strategies:

$$\boldsymbol{\alpha} = \begin{bmatrix} 1-\sigma(a) \\ \sigma(a) \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} 1-\sigma(b) \\ \sigma(b) \end{bmatrix}$$

- Alice's total expected reward is

$$E[r_A] = \big(1-\sigma(a)\big)E[r_A|A=0] + \sigma(a)E[r_A|A=c] = \boldsymbol{\alpha}^T \boldsymbol{A} \boldsymbol{\beta}$$

- Can Alice adjust $\sigma(a)$, using machine learning, to maximize $E[r_A]$?
- Bob's total expected reward is

$$E[r_B] = E[r_B|B=0]\big(1-\sigma(b)\big) + E[r_B|B=1]\sigma(b) = \boldsymbol{\alpha}^T \boldsymbol{B} \boldsymbol{\beta}$$

- Can Bob adjust $\sigma(b)$, using machine learning, to maximize $E[r_B]$?

Simultaneous Gradient Ascent of Two Criteria

- Suppose both Alice and Bob are using mixed strategies:

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 - \sigma(a) \\ \sigma(a) \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} 1 - \sigma(b) \\ \sigma(b) \end{bmatrix}$$

- On successive days, they each try to improve their strategies using gradient ascent:

$$\begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \begin{bmatrix} a \\ b \end{bmatrix} + \eta \begin{bmatrix} \frac{\partial E[r_A]}{\partial a} \\ \frac{\partial E[r_B]}{\partial b} \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \eta \begin{bmatrix} \left(\frac{\partial \boldsymbol{\alpha}}{\partial a}\right)^T \boldsymbol{A}\boldsymbol{\beta} \\ \boldsymbol{\alpha}^T \boldsymbol{B} \frac{\partial \boldsymbol{\beta}}{\partial b} \end{bmatrix}$$

Nash Equilibrium is an Equilibrium

If we start at the equilibrium, $\boldsymbol{\alpha}^T = \left[\frac{1}{2}, \frac{1}{2}\right], \boldsymbol{\beta}^T = \left[\frac{1}{2}, \frac{1}{2}\right]$, what happens?

$$\frac{\partial E[r_A]}{\partial a} = \left(\frac{\partial \boldsymbol{\alpha}}{\partial a}\right)^T \boldsymbol{A\beta} = \begin{bmatrix} 1 - \frac{\partial \sigma(a)}{\partial a} \\ \frac{\partial \sigma(a)}{\partial a} \end{bmatrix}^T \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 - \sigma(b) \\ \sigma(b) \end{bmatrix}$$

$$= \left[\frac{3}{4}, \frac{1}{4}\right] \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = 0$$

Derivative is zero! $a \leftarrow a + \eta \frac{\partial E[r_A]}{\partial a} = a$! The gradient update does not change the agent's strategy, i.e., this is an equilibrium.

…but it's an Unstable Equilibrium!!

If we start off the equilibrium, $\boldsymbol{\alpha}^T = \left[\frac{1}{2}, \frac{1}{2}\right], \boldsymbol{\beta}^T = \left[\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon\right]$, what happens?

$$\frac{\partial E[r_A]}{\partial a} = \left(\frac{\partial \boldsymbol{\alpha}}{\partial a}\right)^T \boldsymbol{A\beta} = \begin{bmatrix} 1 - \frac{\partial \sigma(a)}{\partial a} \\ \frac{\partial \sigma(a)}{\partial a} \end{bmatrix}^T \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 - \sigma(b) \\ \sigma(b) \end{bmatrix}$$

$$= \left[\frac{3}{4}, \frac{1}{4}\right] \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} - \epsilon \\ \frac{1}{2} + \epsilon \end{bmatrix} = +\epsilon$$

Derivative is positive! $a \leftarrow a + \eta \frac{\partial E[r_A]}{\partial a} = a + \epsilon!$ Bob starts away from equilibrium $\Longrightarrow$ Alice is driven to also leave the equilibrium

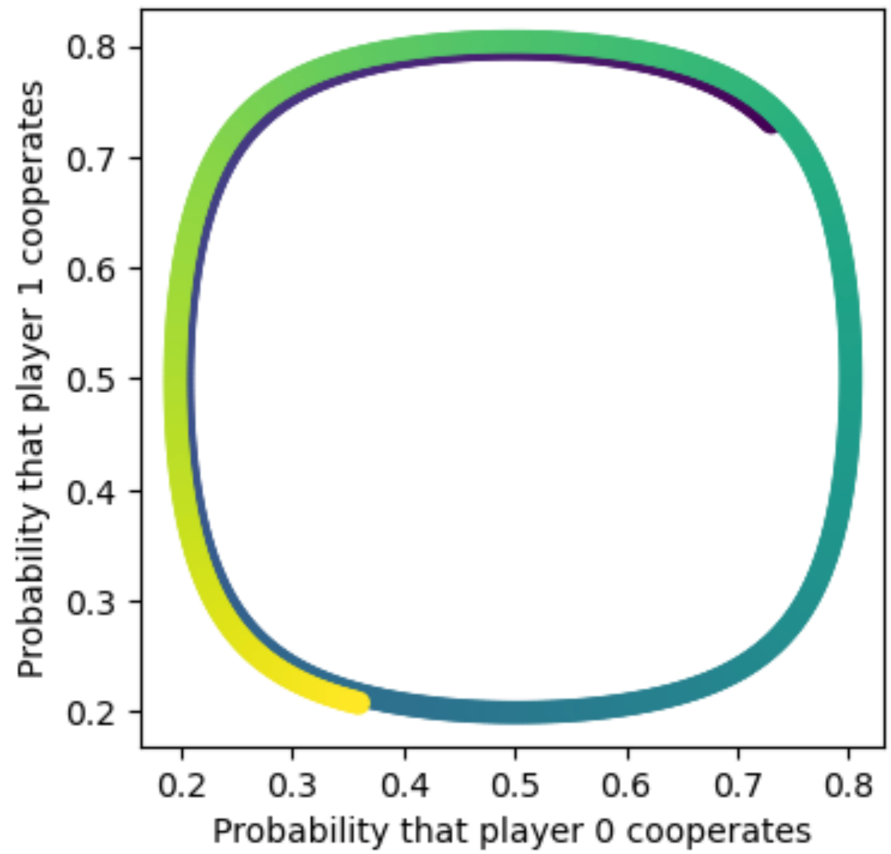
Simultaneous gradient ascent does not always converge to the Nash equlibrium

- Unlike gradient ascent of a single function, simultaneous gradient ascent of two functions is not guaranteed to converge!!
- The graph at right is the sequence of vectors

$$\begin{bmatrix} P(A=c) \\ P(B=c) \end{bmatrix} = \begin{bmatrix} 1/(1+e^{-a}) \\ 1/(1+e^{-b}) \end{bmatrix}$$

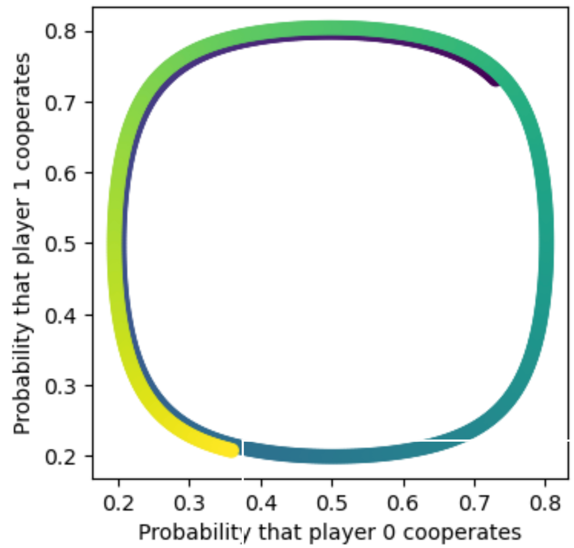
…obtained using

$$\begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \begin{bmatrix} a \\ b \end{bmatrix} + \eta \begin{bmatrix} \frac{\partial E[r_A]}{\partial a} \\ \frac{\partial E[r_B]}{\partial b} \end{bmatrix}$$



Infinite loop

- Why does it never converge?
- If Alice and Bob are in the same location, then Alice goes elsewhere
- If Alice and Bob are in different locations, then Bob follows Alice
- … and so on, forever.
- This is the same loop that we saw last time for the pure strategies. Now we see this loop is also infinite for the mixed strategies!



		Bob Action	
		0	ER
Alice Action	0	-1 / +1	+1 / -1
	1	+1 / -1	-1 / +1

Wait--- Doesn't gradient ascent converge?

- Yes. Gradient ascent converges (if the step size is small enough and there is a local optimum). But gradient ascent means that both a and b are chasing after the ***SAME GOAL***. For example, if they're both trying to improve Alice's day, then the result would converge:

$$\begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \begin{bmatrix} a \\ b \end{bmatrix} + \eta \begin{bmatrix} \frac{\partial E[r_A]}{\partial a} \\ \frac{\partial E[r_A]}{\partial b} \end{bmatrix}$$

- …but in "Simultaneous Gradient Ascent," a and b are trying to optimize ***DIFFERENT GOALS***, and their goals might be incompatible:

$$\begin{bmatrix} a \\ b \end{bmatrix} \leftarrow \begin{bmatrix} a \\ b \end{bmatrix} + \eta \begin{bmatrix} \frac{\partial E[r_A]}{\partial a} \\ \frac{\partial E[r_B]}{\partial b} \end{bmatrix}$$

Does simultaneous gradient ascent always enter an infinite loop?

No, it has various behaviors depending on the game.

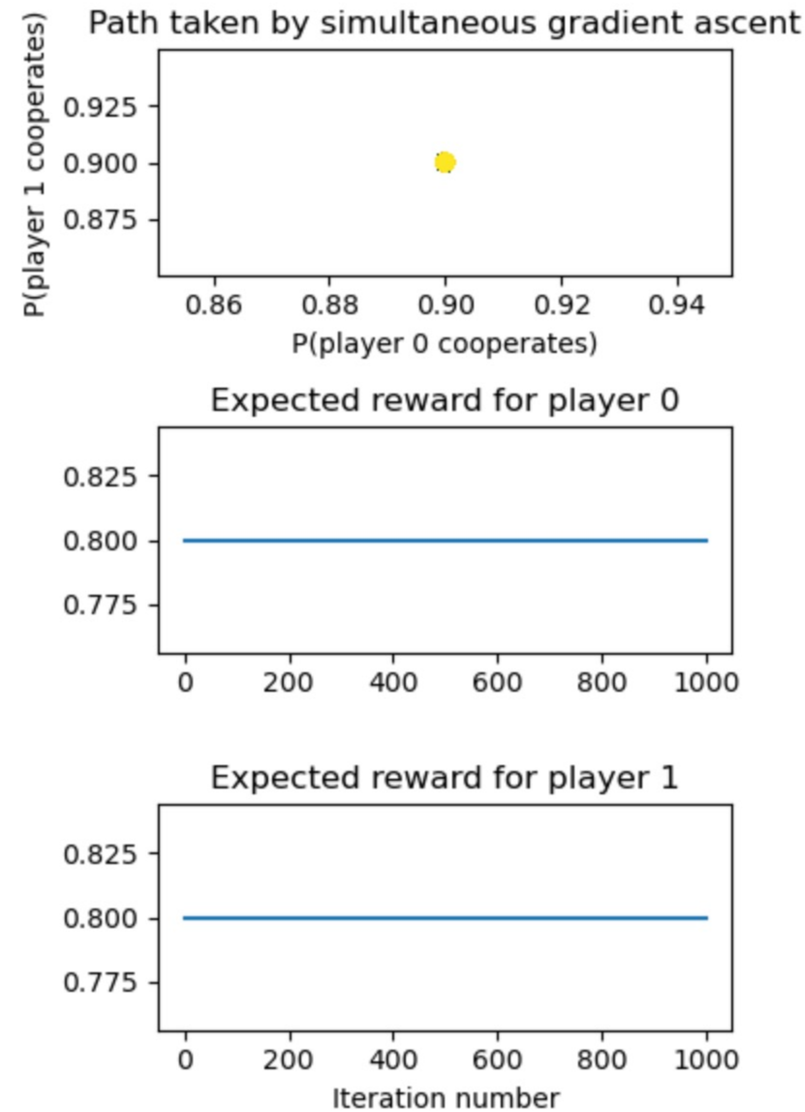
- Paparazzi Game:
  - Start at the equilibrium ⇒ Stays at the equilibrium
  - Start anywhere else ⇒ Enters an infinite loop
- Game of Chicken or Stag Hunt:
  - Start at the mixed equilibrium ⇒ Stays at the equilibrium
  - Start anywhere else ⇒ Converges to one of the two pure-strategy equilibria
- Prisoner's Dilemma (or any game with dominant strategies):
  - Start anywhere ⇒ Converges to the pure-strategy equilibrium

Example: Chicken

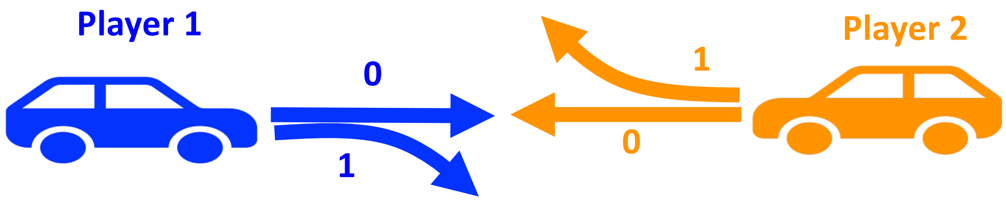
For example, suppose we implement simultaneous gradient ascent for the Game of Chicken:

$$a = a + \eta \frac{\partial E[r_A]}{\partial a}$$
$$b = b + \eta \frac{\partial E[r_B]}{\partial b}$$

- The mixed equilibrium is the pair of strategies $(\sigma(a) = 0.9, \sigma(b) = 0.9)$ where the gradients are zero, so neither player has any rational motivation to change strategies.



Example: Chicken



Bob \ Alice	0	1
0	Alice: -10, Bob: -10	Alice: -1, Bob: 2
1	Alice: 2, Bob: -1	Alice: 1, Bob: 1

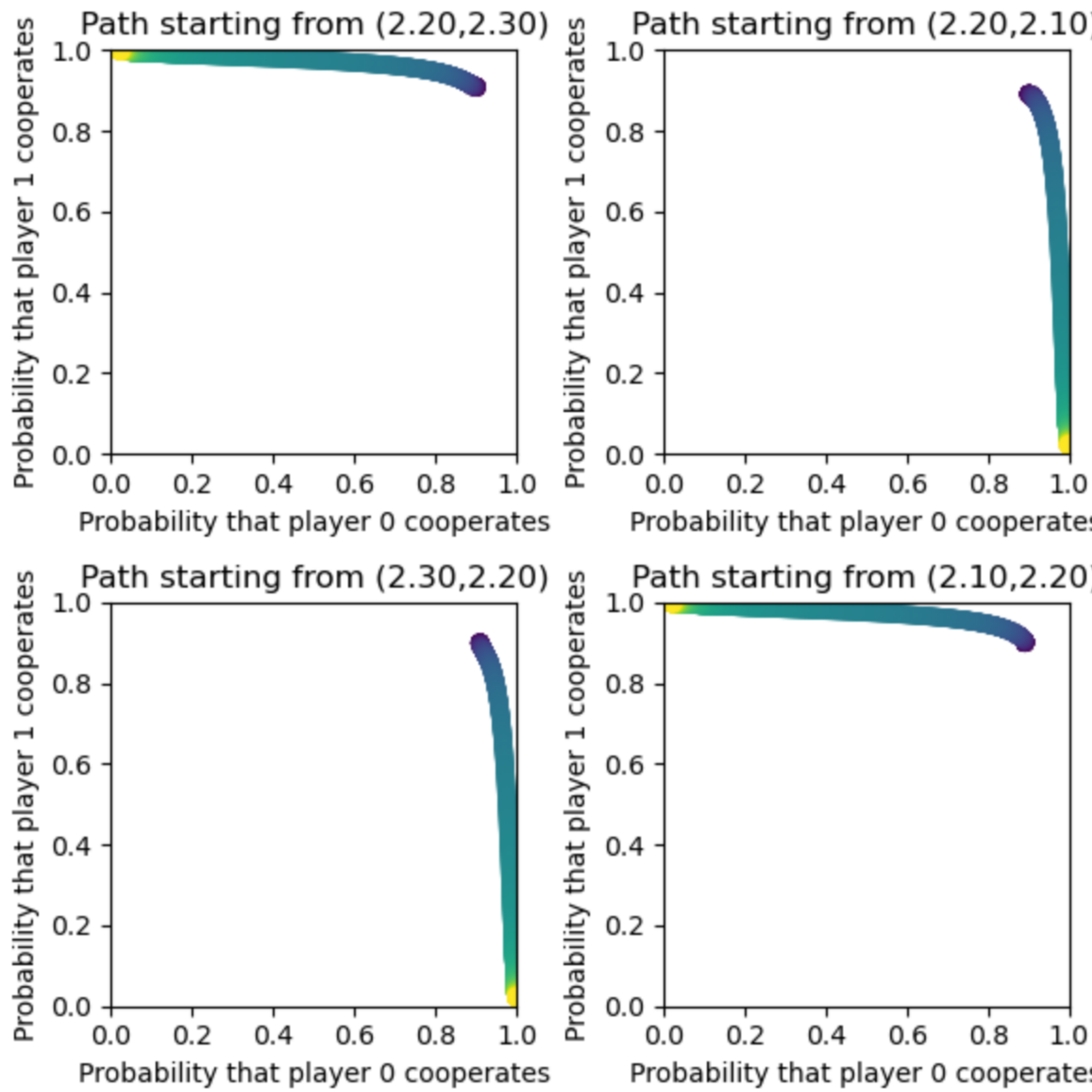
The mixed equilibrium is unstable because a small change in the equilibrium results in a situation that causes the players to move even farther away from equilibrium. For example, suppose that Alice decides to cooperate less often, $P(A=1)=\frac{8}{10}$ instead of $\frac{9}{10}$. Then

$$\frac{\partial E[r_B]}{\partial z_B} = \boldsymbol{p}_A^T \boldsymbol{R}_B \frac{\partial \boldsymbol{p}_B}{\partial z_B} = \left[\frac{2}{10}, \frac{8}{10}\right] \begin{bmatrix} -10 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{9}{100} \\ \frac{9}{100} \end{bmatrix} = +\frac{18}{1000}$$

Since $\frac{\partial E[r_B]}{\partial z_B}$ is positive, it is rational for Bob to increase $P(B=1)$. In response, Alice further decreases $P(A=1)$, until eventually $P(B=1)=1$ and $P(A=1)=0$.

Why is it unstable?

- If Alice cooperates with probability even slightly more than 0.9, then Bob gets better reward by always defecting -> converge to the (C,D) equilibrium.
- If Alice cooperates with probability even slightly less than 0.9, then Bob gets better reward by always cooperating -> converge to the (D,C) equilibrium.



Outline

- The Game of Chicken
- Mixed equilibrium: Randomness is rational behavior
- Adversarial learning
- Generative adversarial networks

Unsupervised learning

Given $\mathcal{D} = \{x_1, \ldots, x_n\}$, learn G so that $P(G = x) \approx P(X = x)$.

- Maximum likelihood: unseen cases have probability zero

$$P(G = x) = \frac{\#\text{ times } x \text{ occurs in } \mathcal{D}}{n}$$

- Laplace smoothing: unseen cases all have the same probability

$$P(G) = \frac{k + \#\text{ times } x \text{ occurs in } \mathcal{D}}{\sum_{x \in \mathcal{X}}(k + \#\text{ times } x \text{ occurs in } \mathcal{D})}$$

Unsupervised learning

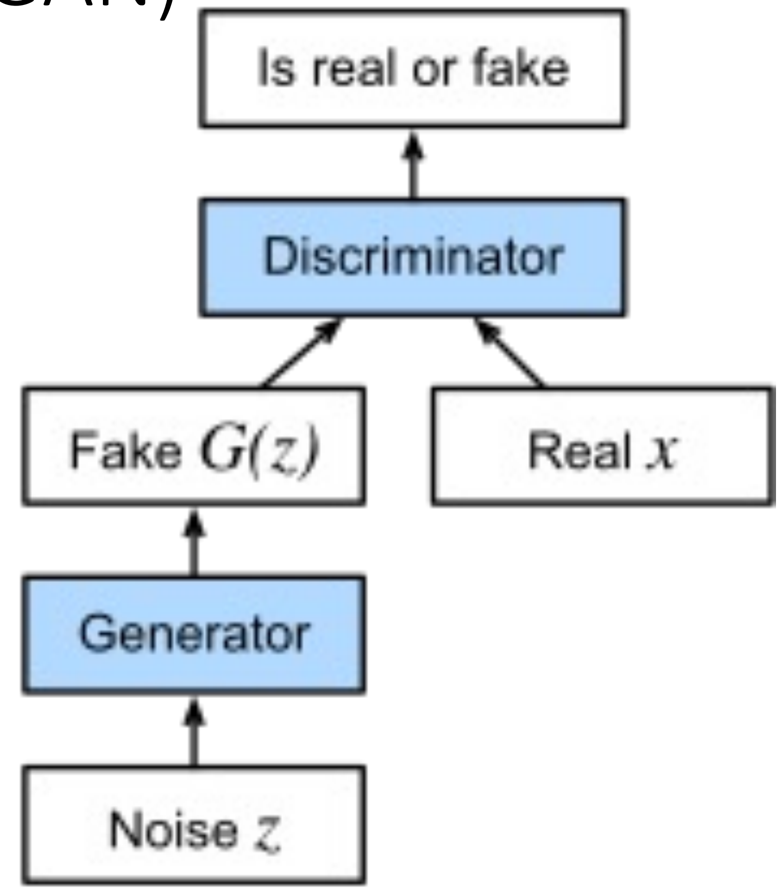
Neither maximum likelihood nor Laplace smoothing is very good for complex random variables. For example, suppose $\mathcal{X}$ is the set of all face images, and we want to train a neural network G so that $P(G = x) \approx P(X = x)$. We would prefer a network to generate images like the one on left, not the one on right:

https://commons.wikimedia.org/wiki/File:Outdoors-man-portrait_(cropped).jpg

https://en.wikipedia.org/wiki/File:Pablo_Picasso,_1910,_Woman_with_Mustard_Pot_(La_Femme_au_pot_de_moutarde),_oil_on_canvas,_73_x_60_cm,_Gemeentemuseum,_The_Hague._Exhibited_at_the_Armory_Show,_New_York,_Chicago,_Boston_1913.jpg

Generative adversarial network (GAN)

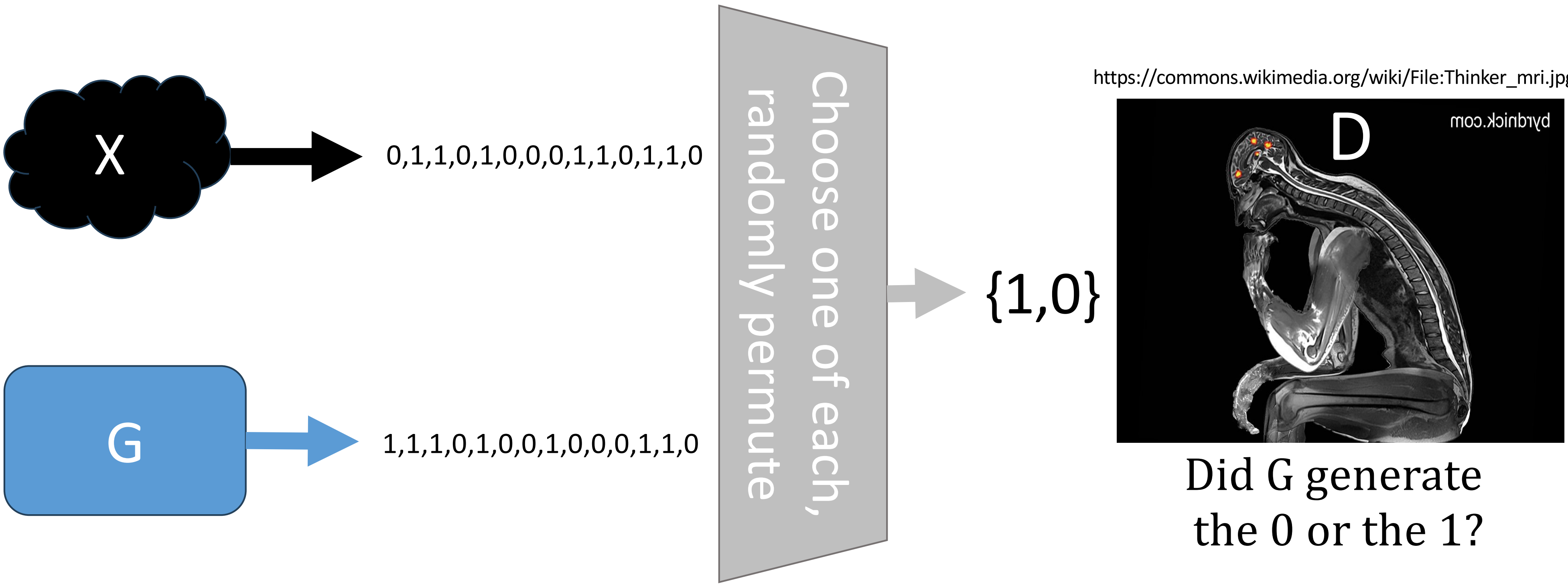
- A generative adversarial network is composed of two networks, a generator (G) and a discriminator (D)
- The generator is trained so that $P(G = x) \approx P(X = x)$, where X is some type of data in the real world
- The discriminator tries to tell the difference between G and X
- If the discriminator can tell the difference, then the discriminator wins
- If the discriminator can't tell the difference, then the generator wins



https://commons.wikimedia.org/wiki/File:Generative_adversarial_network.svg

GAN as a game

- X is a random bit
- G must generate one bit without seeing X
- D gets to see X and G, and needs to decide which one is G



GAN as a game

- If X and G are the same, all rewards are zero.
- If X and G differ, and D can tell which one is G, then D gets rewarded, G gets penalized.
- If X and G differ, and D is incorrect, then D gets penalized, and G gets rewarded.

		Generator	
	X=0	G=0	G=1
Discriminator	D=0	0, 0	-1, 1
	D=1	0, 0	1, -1

		Generator	
	X=1	G=0	G=1
Discriminator	D=0	1, -1	0, 0
	D=1	-1, 1	0, 0

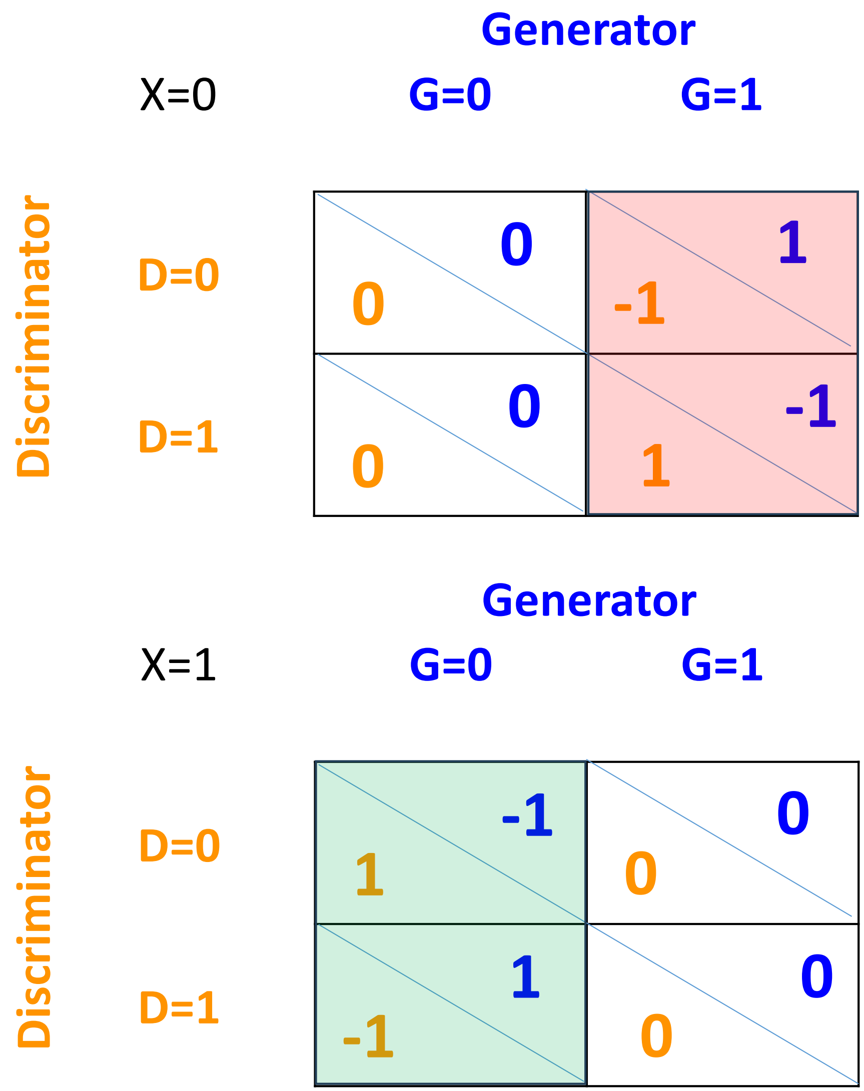
Outcome probabilities

Suppose, independent of one another,

- $X = 1$ with probability P_X
- $G = 1$ with probability P_G

The rewards are all based on the difference between the probabilities of these two rectangles:

$$P(X=0, G=1) - P(X=1, G=0)$$
$$= P_G(1-P_X) - P_X(1-P_G)$$
$$= P_G - P_X$$



Expected rewards

If the discriminator chooses to say that $D = 0$ is the truth, the expected rewards are

$$E[r_D(X, G, D = 0)]$$
$$= P(X = 1, G = 0) - P(X = 0, G = 1)$$
$$= P_X - P_G$$
$$E[r_G(X, G, D = 0)] = -(P_X - P_G)$$

If the discriminator chooses to say that $D = 1$ is the truth, the expected rewards are

$$E[r_D(X, G, D = 1)] = P_G - P_X$$
$$E[r_G(X, G, D = 1)] = -(P_G - P_X)$$

X=0		Generator G=0	Generator G=1
Discriminator	D=0	0 | 0	-1 | 1
Discriminator	D=1	0 | 0	1 | -1

X=1		Generator G=0	Generator G=1
Discriminator	D=0	1 | -1	0 | 0
Discriminator	D=1	-1 | 1	0 | 0

Rational behavior

The **discriminator** should maximize its expected reward, so it should always choose:

- Always choose $D = 1$ if $P_G > P_X$
- Always choose $D = 0$ if $P_G < P_X$
- Choose with 50/50 probability if $P_G = P_X$

The **generator** should maximize its expected reward, so it should choose:

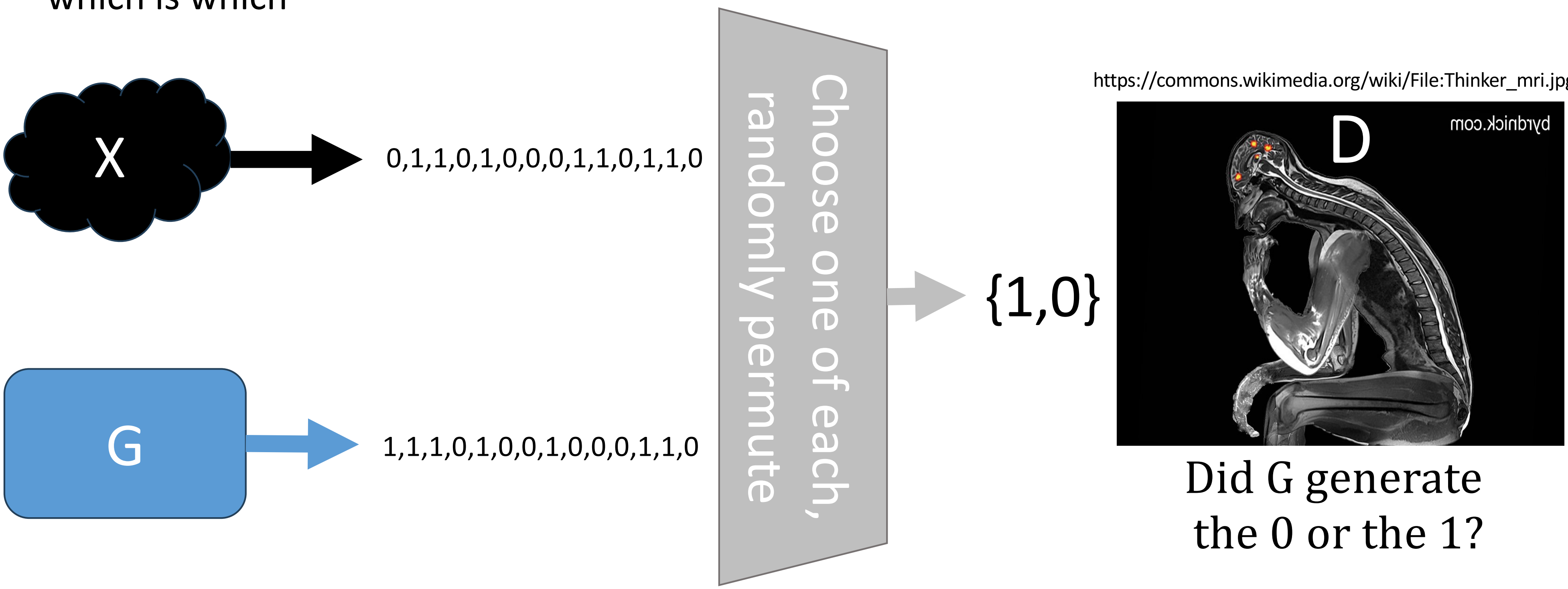
- Always generate $G = 1$ if $P(D = 0) > 0.5$
- Always generate $G = 0$ if $P(D = 1) > 0.5$
- Generate with exactly $P_G = P_X$ if $P(D = 1) = 0.5$

X=0 (Discriminator \ Generator)	G=0	G=1
D=0	0 / 0	-1 / 1
D=1	0 / 0	1 / -1

X=1 (Discriminator \ Generator)	G=0	G=1
D=0	1 / -1	0 / 0
D=1	-1 / 1	0 / 0

Nash equilibrium: Given the other player’s behavior, neither player has a reason to change their strategy.

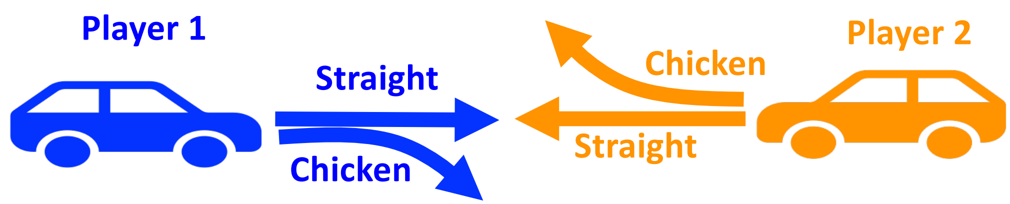
- The generator tries to match the data distribution as exactly as possible
- The discriminator has no choice but to choose uniformly at random, since it doesn’t know which is which



GAN: Unstable Nash equilibrium

- Notice that gains for the GAN are asymmetric:
  - Whenever the generator wins, the discriminator loses
  - Whenever the discriminator wins, the generator loses
- For this reason, the equilibrium is unstable, just like the paparazzi game! --- GANs can be very hard to train
- Some possibilities:
  - After every update of the generator, train the discriminator to fully converge, so that the discriminator tells the generator exactly how it should update next
  - Add extra terms to the loss functions to help convergence (one method, called “symplectic loss,” is modeled after the dynamics of a decaying satellite orbit)

Summary



		Alice: Defect	Alice: Cooperate
Bob	Defect	Alice -10, Bob -10	Alice -1, Bob 2
Bob	Cooperate	Alice 2, Bob -1	Alice 1, Bob 1

- The Game of Chicken
- Mixed equilibrium: Randomness is rational behavior because

$$\begin{bmatrix}-1\\1\end{bmatrix}^T \boldsymbol{A\beta} = \mathbf{0}. \text{ and. } \boldsymbol{\alpha}^T \boldsymbol{B} \begin{bmatrix}-1\\1\end{bmatrix} = 0$$

- Adversarial learning (doesn't always converge)

$$a = a + \eta \frac{\partial E[r_A]}{\partial a}. \text{ and. } b = b + \eta \frac{\partial E[r_B]}{\partial b}$$

- Generative adversarial networks